Significance of the chromosomal positions of the genes encoding Integration Host Factor and DNA gyrase in *Salmonella.*

Thesis presented for the degree of Doctor of Philosophy

Trinity College Dublin

2021

By German Pozdeev

Declarations

I declare that this thesis has not been submitted as an exercise for a degree at this or any other university and it is entirely my own work, with the following exception:

Bioinformatic analyses of whole genome sequencing was performed by Dr. Aalap Mogre.

I agree to deposit this thesis in the University's open access institutional repository or allow the library to do so on my behalf, subject to Irish Copyright Legislation and Trinity College Library conditions of use and acknowledgement.

_____

German Pozdeev

Summary

Multiple studies paved the way to realisation of the importance of gene arrangement along a bacterial chromosome. Patterns have been observed between the position of the genes that encode global regulators, their timely expression during the growth cycle, proximity to their targets and gene orientation with the consequences on local DNA topology. Further studies sought to intervene with the existing chromosomal gene order by altering positions of promoters, open reading frames and entire genes. This approach proved useful to deeper understand the significance of the chromosomal positions of the selected genes, the outcomes of breaching such and the general biology of genes, proteins that they encode and chromosomal dynamics. This knowledge is essential for creation of artificial biological systems or even entire organisms that can be engineered to fulfil a desired task in a fully predictable and controlled way.

In this thesis, two studies are described that built upon the idea of the chromosomal gene importance, shared methodology and yielded results that convey similar message, yet with unique highlights. In both studies *Salmonella* was used as a model organism. In the first study, *ihfA* and *ihfB* – genes that encode nucleoid-associated protein IHF, have been repositioned in several ways. A reciprocal exchange of the open reading frames (ORFs) of the two genes resulted in the OrfSwap$^{ihfA\text{-}ihfB}$ strain. The ORF of *ihfB* have been moved downstream of *ihfA*, creating the strain with the *ihfAB* operon. The analogous manipulations resulted in the *ihfBA* strain. In the other study, *gyrA* and *gyrB* – genes that encode an essential topoisomerase DNA gyrase, have been arranged in the *gyrBA* operon that naturally exists in many bacteria, other that members of the class gammaproteobacteria. The artificial operonic arrangement was a novel approach to the chromosomal gene rearrangements.

Multiple phenotypic tests, mostly related to growth and morphology of all the strains, highlighted *Salmonella's* ability to successfully tolerate rearrangements of its genes. Minor alterations in the expression of SPI-1 and SPI-2 manifested in the ability of the strains to enter and survive in murine macrophages. Phenotypes, such as motility, competitive fitness, SPI expression and

expression of the repositioned genes, hinted that the strains with the operonic arrangement of the IHF-encoding genes have higher resemblance to the WT than the OrfSwap$^{ihfA\text{-}ihfB}$ strain. The global DNA supercoiling in the *gyrBA* strain was found to be altered in comparison to the WT and tuneable by the growth medium composition, indicating potential broad outcomes on the global gene expression.

The determination of the IhfA and IhfB protein levels by Western blotting in the WT and the OrfSwap$^{ihfA\text{-}ihfB}$ strains led to the discovery that IhfB subunit is present in the WT at a concentration that is almost twice that of the IhfA subunit. This was confirmed by mass spectrometry that also showed that the pattern was swapped in the OrfSwap$^{ihfA\text{-}ihfB}$ strain. It was determined that higher stability of *ihfB* mRNA is responsible for unequal production of IHF subunits. The same MS experiment showed that multiple proteins involved in translation are downregulated in the OrfSwap$^{ihfA\text{-}ihfB}$ strain, leading to the likely decrease in the overall protein production.

Genes that encode DNA gyrase are located separately in gammaproteobacteria but in many other bacteria they are arranged in the operon. This puzzling inconsistency led to the comparative study of all the possible arrangements of these genes among bacteria. Four different groups of *gyrA* and *gyrB* arrangements were found that are based on 1 – operonic or separate positioning, 2 – conservation of the genetic environment 5' of *gyrB*. It was found that *gyrB* position is more conserved than that of *gyrA*.

Overall, this work confirmed the exceptional robustness of *Salmonella* towards chromosomal repositions; advanced the understanding of differential positioning of *gyrA* and *gyrB* between bacterial taxa, provided compelling evidence for the exceeding production of IhfB over that of IhfA; and left ample potential for future exploration.

Acknowledgements

No one can do science on their own, the success of this work is thanks to the inhabitants of the Moyne, my friends and family who have been with me through all the ups and downs during these tough but unforgettable four years and to Science Foundation Ireland for funding this project.

First and foremost, I would like to thank my supervisor Charles Dorman for everything, from kindly letting me in his lab for a summer project back in my undergrad, trusting me to do a PhD project under his guidance and until this day. You have been an example of a true scientist for me.

The special thanks to all the Dorman lab members past and present with whom I had a privilege to work. My first microbiology teacher – Aoife, Connie – the invaluable motor of our lab, Aalap – wise and thoughtful, Eoin – bright and encouraging, my scientific sister and brother – Marina and Mick. Mick, your support and mentorship at many levels of my life over the past two years are priceless and I'm grateful to you, brother.

Thank you to all the people who comprise the soul of the Moyne. Every generation here have been amazing. From the now-legendary Dara, Chandie. Kieran, Stefani, Aisling, Michelle, Isa, Niamh; the current lovely squad Amy, Dani, Brenda, Marty, Mohammed, Mary, Rob; and to the future leaders Candida, Fergal, Nicole; guys you made all the difference and your friendships are so important for me. Despite people come and go, the succession of the traditions in this family home is unbelievable! I hope that the 2020 won't change anything. Long live the Moyne!

A huge thank you to everyone in the "Prep room" and the office. Stephen and Ronan, thanks for the most human and sincere chats that we had in the "Bear cave". Jayne, without you everything would break apart.

Thank you to Eugene Dillon of UCD Conway Mass Spectrometry Core for his invaluable advices on initial steps of the MS analysis.

Thank you to my committee members Prof. Carsten Kröger and Prof. Joan Geoghegan for your wise and stress-free guidance throughout these years.

Thank you to my parents who always believed in me and influenced my decision to dedicate myself to research. I am standing where I am because of you. Thank you to all my extended family in Russia for your prayers, encouragements and long-standing belief in me.

Thank you to my wife Ira who patiently waited for me during the long nights in the lab, made our little home a quiet harbour and inspired me every day. I love you endlessly and look forward to dedicating all my time to you and to our future adventures!

Publications and Conference Communications

Publications:

Pozdeev, G., Mogre, A., Dorman, C. J. (2021), Consequences of producing DNA gyrase from a synthetic *gyrBA* operon in *Salmonella enterica* serovar Typhimurium. *Mol Microbiol*, in preparation

Pozdeev, G., Mogre, A., Beckett, M. C., Dorman, C. J. (2021), in preparation

Conferences Attended:

German Pozdeev, Charles J. Dorman. Significance of the chromosomal positions of the genes encoding DNA gyrase in *Salmonella enterica* serovar Typhimurium. Poster presentation at Microbes in Medicine: A Century of Microbiology at Trinity College Dublin October 2019. Dublin, Ireland.

German Pozdeev, Charles J. Dorman. Significance of the chromosomal positions of the genes encoding nucleoid-associated protein IHF in *Salmonella enterica* serovar Typhimurium. Poster presentation at Microbiology Society Annual Conference April 2017. Edinburgh, UK.

Table of contents

List of Figures

List of tables

Chapter 1: General introduction

1.1 General introduction

1.1.1 Bacterial genome

At the basis of any living organism lies its genetic code that is inherited from generation to generation and shapes the appearance of the organism, causing it to resemble its relatives via gene expression. The complete set of genes of an organism is referred to as its genome and takes different forms in different organisms. In eukaryotes, the genome is organised within a membrane-enclosed organelle called the nucleus and is composed of pairs of linear chromosomes, number of which can vary greatly depending on a species (Eichler & Sankoff, 2003). Prokaryotic genomes are also diverse and can be composed of chromosomes of different numbers, sizes and shapes, as well as plasmids. The biggest difference between eukaryotic and prokaryotic genomes is the lack of nuclear compartmentalisation in prokaryotes – there is no nuclear membrane. However, this does not mean that prokaryotic genomes are simpler to understand. On the contrary, they are arguably as complex as eukaryotic counterparts due to localisation in a structure referred to as nucleoid, that is composed of DNA and multiple auxiliary proteins (Dame *et al.*, 2020).

The classic bacterial genome is a single circular chromosome, such as the 4.6 Mb chromosome of the model bacterium *Escherichia coli* (Blattner *et al.*, 1997). However, multiple circular chromosomes can be found in other bacteria. Members of *Vibrio* spp. harbour two chromosomes (Heidelberg *et al.*, 2000), while members of *Epulopiscium* spp. contain unprecedented tens of thousands of chromosome copies (Mendell *et al.*, 2008). Linear chromosomes exist in other bacteria such as *Streptomyces* spp. (Tsai *et al.*, 2011) and *Borrelia* spp. The latter has telomeres at both ends of the genome that gives it extra stability (Chaconas & Kobryn, 2010). The plant pathogen *Agrobacterium tumefaciens* harbours one circular and one linear chromosome (Wood *et al.*, 2001). It should be noted that the circular chromosome of *E. coli* can be linearized without a significant effect on the cell cycle or growth (Cui *et al.*, 2007), while the linear chromosome of *Streptomyces* can be circularized, both naturally and artificially, but at a cost to genetic stability (Volff & Altenbuchner, 2000). Thus, the shape

that a bacterial genome acquires helps the organism to maintain genetic information in an optimal state for utilization but is not essential for viability.

Salmonella enterica serovar Typhimurium is the organism that is focus of this work. This bacterium possesses unique characteristics of virulence and is a popular model for study of the intestinal infection in mammals. The genome of *Salmonella* is a single 4.9 Mb circular chromosome and a virulence plasmid. The chromosome, with the genetic features that are important for this work highlighted, is depicted in Fig. 1.1.

Among the features used by bacteria to adapt to the environment is the flagellum – a structure that is assembled through a complex series of events. Rotation of the flagellum allows it to propel the bacterium through its environment in response to chemical gradients through the process known as chemotaxis (Chaban *et al.*, 2015). Another example of environmental adaptation involves the reciprocal production of porin proteins in the outer membrane of *Salmonella* and multiple other Gram-negative bacteria. Porin OmpC is preferentially expressed in high osmolarity, while porin OmpF is expressed at low osmolarity and has a pore size that can allow passage of larger solutes (Nikaido, 2003). Their appropriate expression in response to environmental cues ensures the optimal performance of the bacterium in the context of its surroundings. The production by *Salmonella* of specialised virulence structures, such as type 3 secretion machinery, is a further example of adaptation to the environment, in this case to the small intestine of a mammalian host. The genes that encode the type 3 secretion systems of *Salmonella* are clustered in pathogenicity islands on the chromosome – large contiguous tracts of genes that have been acquired by horizontal gene transfer. Their colocation within the islands facilitates the coregulation of the expression of these genes in complex and coordinated control networks that are reminiscent of those governing the expression of the flagellum genes (Schmidt & Hensel, 2004).

The colocation of genes encoding components of a common macromolecular system is intuitively obvious. However, some other features related to the genomic composition are less easily understood. While these features are

stably inherited and therefore maintained by the natural selection, the immediate reason behind their presence is often obscure. For example, the reason for the separate chromosomal location of some genes that code for subunits of the same proteins such as DNA gyrase, topoisomerase IV, IHF, HU is not immediately obvious. *A priori*, the co-expression and co-regulation of such genes would seem to be more advantageous. As bacteria are frequent subjects of horizontal gene transfer, the content of their genomes can vary in ways that differ from the genome patterns seen in eukaryotes that arise from patterns of microevolution that can be hard to notice and study (Bobay & Ochman, 2017). How important is gene position within the chromosome? This question is central to the scope of this work. A thorough understanding of the aspects of the chromosomal gene positioning, the nucleoid, and *Salmonella* infection is required to appreciate the aims and results of this work. These topics will be introduced in this chapter.

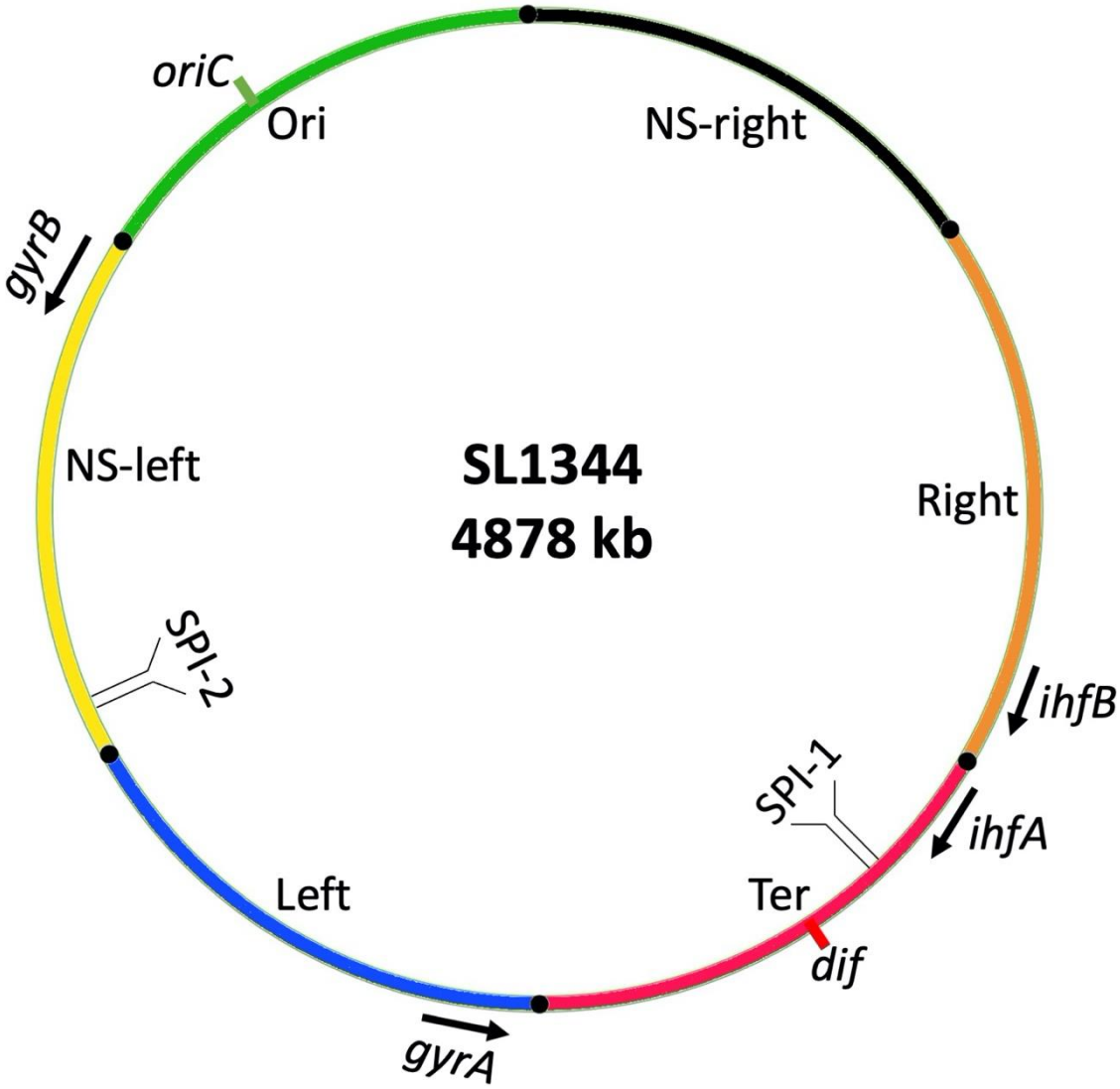


Figure 1.1 Circular chromosome of *Salmonella enterica* serovar Typhimurium str. SL1344. Macrodomains are highlighted with colours. Positions of *oriC* and dif are indicated. Positions and orientations of *ihfA*, *ihfB*, *gyrA* and *gyrB*, as well as *Salmonella* pathogenicity islands SPI-1 and SPI-2 are indicated. The schematic drawing is not to scale.

1.1.2 The nucleoid

In spite of its seeming simplicity, the bacterial chromosome is a complex structure. Figure 1.1 depicts the chromosome as a circle, in reality it is a multi-layered, tightly folded mass of DNA that is highly structured at the nanometre, micrometre and millimetre scale with the aid of DNA-binding proteins, RNA, and DNA supercoiling. This allows genetic material of the chromosome to be accessed when needed by the transcriptional and replication machinery, and yet to be able to fit inside a cell that has length of only 1-2 μm, a diameter of 1 μm and a volume of 1.5 fl (Volkmer & Heinemann, 2011).

At the coarsest level of organisation, the chromosome is subdivided into four macrodomains – the regions that do not interact with each other (Valens *et al.*, 2004) and two non-structured regions – the regions that can interact with the adjacent macrodomains (Fig. 1.1). The macrodomains occupy distinct positions inside the cell throughout the cell cycle (Fig. 1.2a). This localisation is important for correct chromosome segregation and is controlled by macrodomain-specific proteins. Ter-binding MatP ensures timely segregation (Mercier *et al.*, 2008), SlmA prevents FtsZ ring formation except at the Ter (Monterroso *et al.*, 2019), SeqA prevents over-initiation of replication at the *oriC* (Lu *et al.*, 1994) and takes part in gene regulation and chromosome organisation by binding at sites other than *oriC* (Waldminghaus & Skarstad, 2009). The major player in organising this chromosomal localization in *Salmonella* is the MukBEF protein complex that belongs to the structure maintenance chromosome (SMC) family of proteins and binds throughout the genome but with a preference to the Ori macrodomain. MukBEF serves as a scaffold to navigate the replicated origin towards the pole of a cell (Danilova *et al.*, 2007). According to one of the models of action, it does so via series of ATP binding/ hydrolysis that causes the MukBEF complex to “walk” along the chromosome (Badrinarayanan *et al.*, 2012).

Further compaction of the nucleoid is achieved by small abundant proteins called nucleoid-associated proteins (NAPs) and DNA supercoiling. Histone-like nucleoid-structuring protein H-NS is a NAP, dimers of which are known to bind AT-rich DNA to carry out gene regulatory and architectural functions (Gordon *et*

al., 2011). H-NS is able to form two types of nucleoprotein complexes: nucleoprotein filaments and DNA-DNA bridges (Qin *et al.*, 2019). These can be interconverted with the aid of conformation change of H-NS dimers (van der Valk *et al.*, 2017; Shahul Hameed *et al.*, 2019). The H-NS DNA bridging activity shapes boundaries for nucleoid microdomains (Hardy & Cozzarelli, 2005).

Other NAPs important for nucleoid compaction are DNA bending proteins Integration Host Factor (IHF), HU, Factor for Inversion Stimulation (FIS), DNA protection against starvation (Dps) and curved DNA-binding protein A (CbpA). IHF and HU are paralogous proteins that are very similar in structure, but distinct in binding preferences (Swinger & Rice, 2004). The outcome of their binding is substantial bending of DNA that is able to alter its path, thereby contributing to altering local chromosomal architecture. FIS, like H-NS, can bring together genome-wide FIS-binding sites, organising branched plectonemes of DNA (Schneider *et al.*, 2001). Binding of Dps and CbpA to DNA at times of stress and starvation condenses the chromosome (Meyer & Grainger, 2013). This condenses DNA to the state of crystalline structures that ensure DNA protection but allow all cellular processes to occur due to formation of phase-separated organelles (Janissen *et al.*, 2018).

DNA supercoiling is a process whereby DNA is under- or over-wound and reflects the extent of molecular interactions in which DNA is involved and is a manifestation of stored potential energy that can be used to drive DNA transactions, such as transcription and replication, that rely on DNA strand separation. The extent of DNA compaction by supercoiling can change throughout growth in batch culture. For example, during the period of extensive growth at the exponential phase, DNA is more supercoiled, while during the lag and stationary phases of growth it is more relaxed (Cameron & Dorman, 2012) (Fig. 1.2b).

The parameters of DNA supercoiling can be described in terms of the linking number (Lk), that is a number of times one DNA strand passes over the other in a double helix. Lk is composed of the twist (Tw), that is a number of helical turns around the axis of the double helix, and writhe (Wr), that is a number of times the axis of the helix passes over itself (Fig. 1.2c) (Higgins & Vologodskii,

2015). Twist and writhe can interconvert, while the Lk can change only upon breaking of a DNA phosphodiester bond by DNA topoisomerases. Negative supercoiling has the effect of making the DNA molecule more compact. DNA free of any protein possesses the superhelical density (σ) – the measure of a difference between relaxed and supercoiled states of DNA) of -0.05 (Vologodskii *et al.*, 1979).

Information in DNA can be thought of using the metaphors of digital and analogue information, where the nucleotide sequence approximates to the digital component and the shape, or topology, of the DNA represents the analogue component. DNA shape is determined by the transactions taking place within it (e.g. transcription and replication), the actions of topoisomerases and the binding activities of proteins, including NAPs. The AT content of the DNA contributes to its intrinsic shape with phased A-tracts, for example, introducing intrinsic bends into the molecule (Haran & Mohanty, 2009; Rohs *et al.*, 2009). Bends and loops in DNA, can physically prevent DNA-protein interactions but can be reversed to promote these interactions, introducing a dynamic dimension to nucleoid architecture. Long-range interactions are facilitated by the nature of nucleoid organisation, allowing sites that are located in separated regions on a chromosome to come together in space (Dame *et al.*, 2020). It is possible that such interactions may facilitate the co-expression of distantly located genes that contribute to a common output and enhance the efficiency of assembly of their products into macromolecular assemblies. As NAPs and DNA supercoiling are two of the factors responsible for nucleoid structuring and maintenance, the particulars of the NAP IHF and the topoisomerase DNA gyrase – important members of these nucleoid-structuring categories, as well as other chosen players which relate to the current work, will be discussed below.

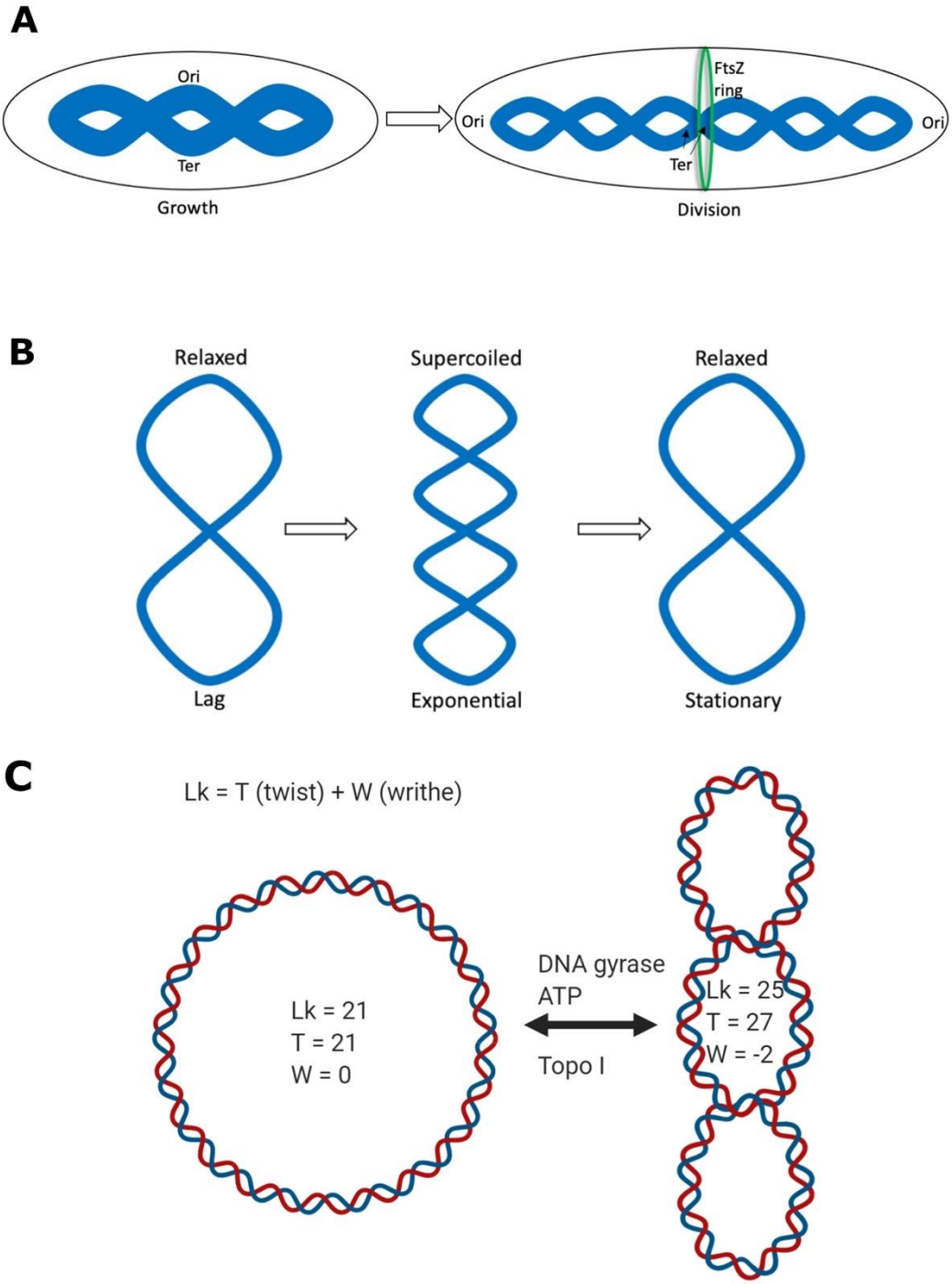


Figure. 1.2 Nucleoid is a dynamic entity. a) Ori and Ter macrodomain positioning relative to the middle and the poles of a cell before and during division. b) DNA is more supercoiled at the exponential phase of growth than at the lag phase and the stationary phase. c) DNA can change its linking number (Lk) via the change in twist (Tw) and/or writhe (Wr). Based on (Dorman, 2013; Dame, 2020).

1.1.3 Integration Host Factor (IHF) and other NAPs

IHF was discovered as a co-factor important for the site-specific recombination of bacteriophage lambda into and excision from the chromosome of *E. coli* (Miller & Friedman, 1980), hence the name. IHF is a heterodimer of IhfA and IhfB subunits, although evidence is accumulating that homodimeric forms of IHF also exist and play distinct biological functions (Mangan *et al.*, 2006; Prieto *et al.*, 2012). Every function of IHF is mediated by its ability to sharply bend DNA to angles of up to 180 degrees. Each subunit consists of a β-pleated sheet arm that mediates DNA-binding and a body composed of α-helices that promotes dimerization (Rice *et al.*, 1996). IHF binds specifically to a conserved sequence 5'-WATCAANNNNTTR-3', where W is A or T, R is G or A and N is any nucleotide (Goodrich *et al.*, 1990). This consensus sequence consists of two conserved patches – before and after NNNN. The first patch is bound by the arm of IhfA in the minor groove of DNA, while the second is bound by the body of IhfB (Fig. 1.3a). Intercalation of prolines that are present in arms of both subunits into DNA produces two large kinks in DNA making the DNA execute a U-turn, wrapping around the protein in a dynamic nucleoprotein interaction (Connolly *et al.*, 2018). IHF binds to its consensus binding sequence strongly (K_d = 2-20 nM), while binding to non-specific sequences is weak (Swinger & Rice, 2004; Henneman *et al.*, 2018). Prior to recognising its consensus binding site, IHF scans DNA in a series of non-specific binding events that result in partial DNA bending (Velmurugu *et al.*, 2018).

IHF can play both positive and negative regulatory role in the initiation of chromosome replication. Binding of DnaA in complex with ATP to its binding sequences at *oriC* (DnaA boxes) initiates DNA replication via oligomerisation and unwinding of the DNA helix at the AT-rich DNA unwinding element (DUE). Binding of IHF to its consensus sequence at *oriC* stimulates binding and oligomerisation of DnaA (Grimwade *et al.*, 2000). Binding of IHF at the DARS1 and DARS2 loci promotes synthesis of active DnaA-ATP (Kasho *et al.*, 2014). To inhibit the initiation of replication, IHF binds to the site located away from the *oriC* at the *datA* locus. Binding of IHF there causes sharp DNA bend that in turn

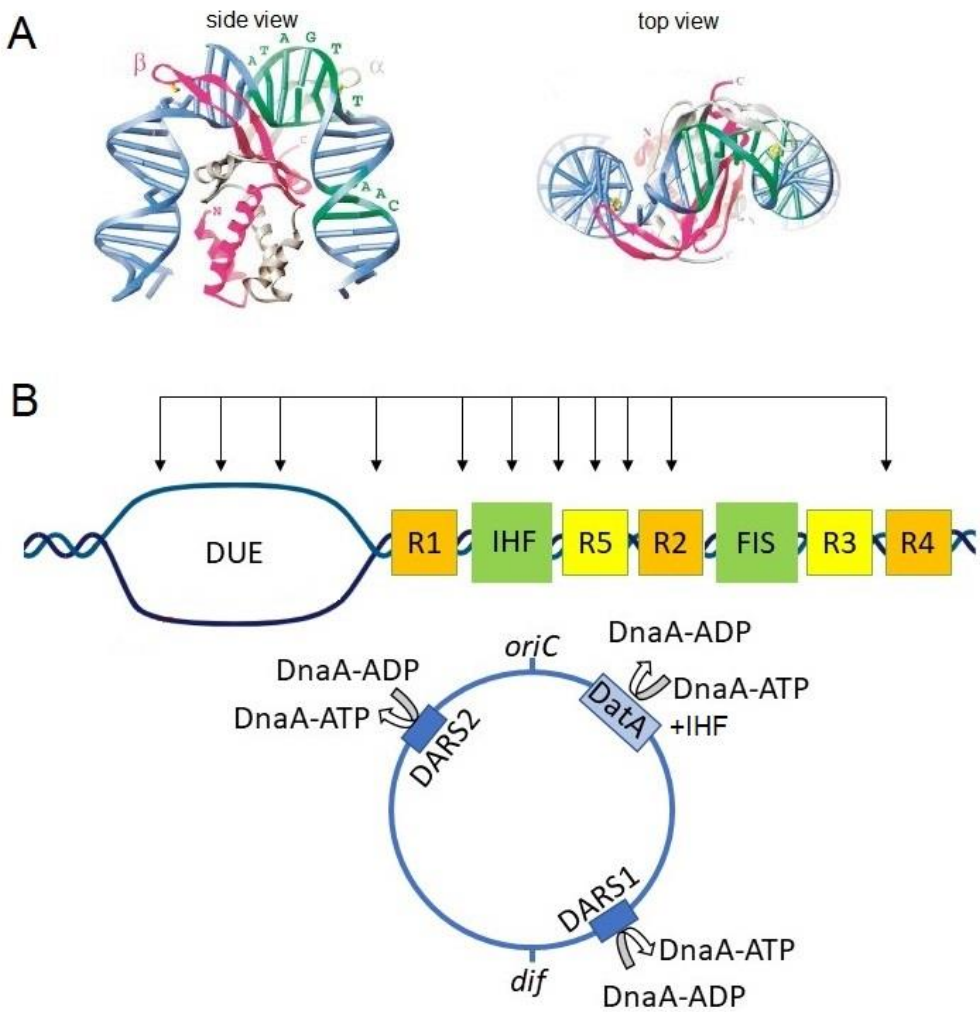


Figure 1.3 IHF structure and function in the initiation of chromosomal replication. a) IHF heterodimer co-crystallised with H' site of bacteriophage lambda. Top and bottom view. b) Top – structure of *oriC*. Initiation of replication begins with IHF binding and introducing a bend in DNA to promote binding of DnaA to DnaA boxes (R1-R5). Orange indicates high-affinity binding sites and yellow indicates low-affinity binding sites. DnaA binding promotes unwinding at the AT-rich DNA Unwinding Element (DUE). Binding of FIS blocks IHF and subsequent DnaA binding, inhibiting the initiation of replication. Binding of SeqA to 11 hemimethylated 5'-GATC-3' sites (black arrows at the top) immediately after the initiation is inhibitory for the immediate re-initiation of chromosomal replication. Bottom – DnaA-ADP to active DnaA-ATP conversion is happening at DARS1 and DARS2 sites. DnaA-ATP to inactive DnaA-ADP conversion occurs at DatA locus upon IHF binding, reducing size of active DnaA pool. Drawn not to scale. Based on (Rice *et al*., 1996; Dorman, 2020).

promotes DnaA binding at this locus, titrating it away from the *oriC* (Kasho et al., 2017). (Fig. 1.3b).

As mentioned above, IHF is required for integration and excision of bacteriophage lambda into/from the *E. coli* genome (Miller & Friedman, 1980; Bushman *et al.*, 1984). Both reactions are catalysed by the bacteriophage-encoded Integrase protein (Int) – a prominent member of site-specific tyrosine recombinase family. Int requires host-encoded IHF, as well as FIS and Xis, in order to bring together distant sites that are required for the recombination (Seah *et al.*, 2014). Another recombination-based system – the *fim* switch, is dependent on IHF as a co-factor. This shows that IHF is not only involved in events within mobile genetic elements such as lambda, but also contributes to regulatory mechanisms in the host genome. The orientation of the invertible element of the *fim* switch determines the appearance of fimbriate or afimbriate cells. Binding of IHF at its two consensus sites within the switch, along with binding of other NAPs and DNA topology, is responsible for the efficient inversion of the switch. IHF bends DNA to allow FimB and FimE tyrosine integrases to carry out the directional recombination reaction between the inverted repeats that surround the invertible element (Dorman & Bogue, 2016) (Fig. 1.4a). The DNA-bending properties of IHF also ensure correct transposition of bacteriophage Mu and transposable elements, such as Tn*10* (Kobryn *et al.*, 2002; Crellin *et al.*, 2004).

IHF is extensively involved in the regulation of transcription. Transcriptomic studies that were using single and double IHF subunit deletion mutants determined the role of IHF in regulation of multiple genes involved in virulence, motility, chemotaxis, stress, metabolism and those encoding stationary-growth-phase-specific (Mangan *et al.*, 2006; Prieto *et al.*, 2012). The mechanisms by which IHF regulates transcription are all dependent on this NAP's DNA bending activity. It can recruit the RpoS sigma factor of RNA polymerase to a promoter, exemplified by the P_{nifH} of *Klebsiella pneumoniae,* the P_{Pu} of the *Pseudomonas putida*, and P_{dps} of *E. coli* (Santero *et al.*, 1992; Macchi *et al.*, 2003; Altuvia *et al.*, 1994). IHF binding can also lead to a displacement of a supercoiling-induced DNA duplex destabilisation (SIDD) structure by transferring energy, that is absorbed by the SIDD in the absence of IHF, from an IHF binding site

downstream to the promoter region, resulting in formation of an open complex at the P_{ilvPG} (Sheridan *et al.*, 1998).

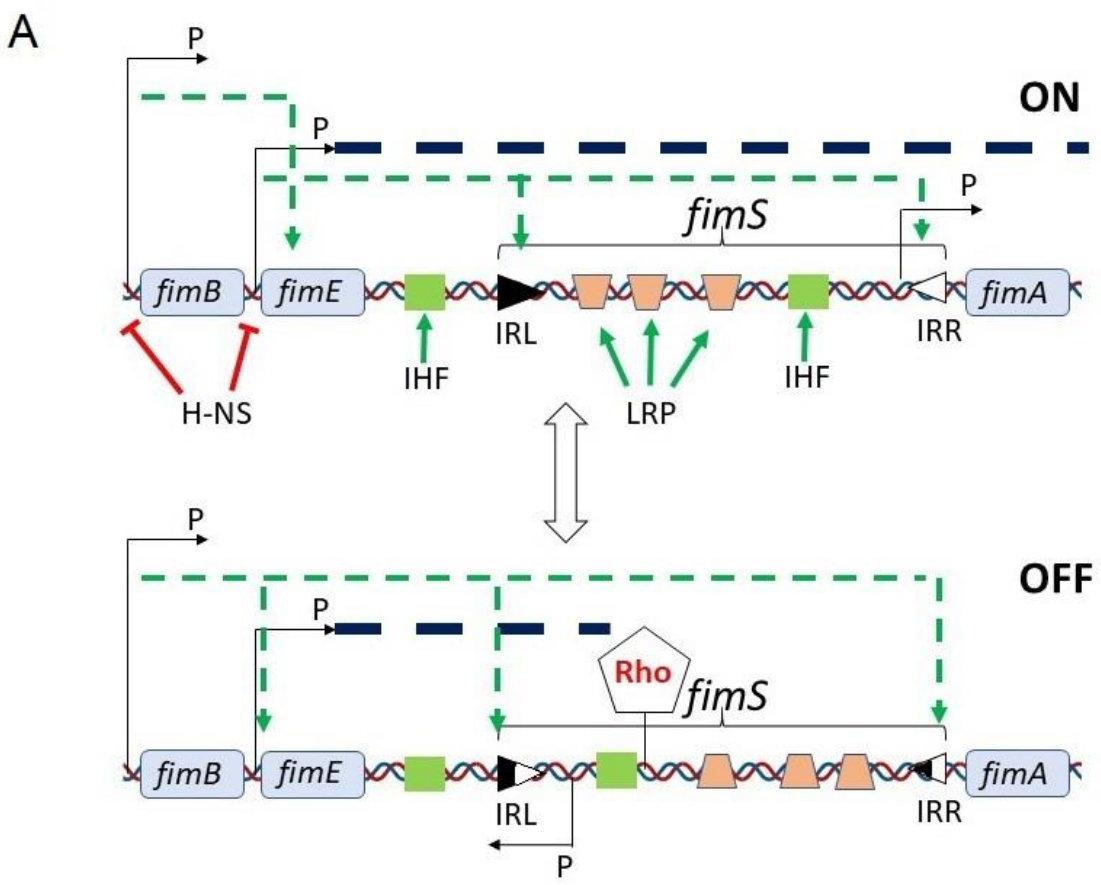
IHF also plays a role in CRISPR-*Cas* immunologic memory. There is an IHF-binding site in the leader peptide of the CRISPR-I locus in *E. coli* (Yosef *et al.*, 2012). Spacer acquisition was inactivated in *E. coli* mutants where IHF mutants was unable to bind to DNA and it was restored when IHF was provided in-*trans* on a plasmid (Nuñez *et al.*, 2016). Targeted deformation of DNA is a prerequisite for the specific recognition of CRISPR leader peptide by Cas1-Cas2 integrase (Fig. 1.4b).

IHF in other bacterial phyla is structurally different but provides similar functionality. IHF of *Streptomyces coelicolor* (sIHF) binds DNA as a monomer and stabilizes, rather than introduces, DNA bending (Nanji *et al.*, 2019). This NAP influences chromosome condensation, segregation, antibiotic production and, thus, gene expression regulation (Swiercz *et al.*, 2013). IHF of *Mycobacterium tuberculosis* (mIHF) is similarly monomeric and structurally unrelated to IHF of *E. coli* and is able to induce DNA bending and compaction (Odermatt *et al.*, 2020). Interestingly, mIHF can complement IHF subunit deletion mutants in *E. coli* for growth against genotoxic stress, nucleoid compaction and recombination (Sharadamma *et al.*, 2014).

It should be noticed that NAP composition between different bacterial phyla differs. For example at least one HU subunit, member of the same IHF/HU fold-containing family, is present in all the bacteria (and other kingdoms of life) (Grove, 2011), while IHF-like proteins are almost exclusively found in *Proteobacteria* (Dey *et al.*, 2017; Kamashev *et al.*, 2017). Even though bacteria in different phyla have unique NAP composition, the general functions of these proteins remain the same – structuring of nucleoid architecture and regulation of gene expression.

The NAP composition of *E. coli* and *Salmonella* is the same. There is a core set of such proteins that are most abundant and are widely recognised as NAPs. Other proteins, such as Hfq, DnaA, MukB, may or may not be referred to as NAPs depending on a context by different researchers (Ali Azam & Ishihama, 2015; Dorman, 2014). There are at least 10 NAPs found in these species (Ali

Azam *et al.*, 1999; Dillon & Dorman, 2010) (Table 1.1). Apart from IHF and HU, some of the most studied NAPs are H-NS, StpA, Dps, Fis, LRP and MukB. Some NAPs are expressed throughout the growth cycle, while others are found only at a particular phase of growth (Ali Azam *et al.*, 1999; Dillon & Dorman, 2010; Ali Azam & Ishihama, 2015). Intrinsic DNA curvature, flexibility and sequence specificity are the parameters that characterise binding sites for most of these proteins (Liu *et al.*, 2018). Post-translational modifications of NAPs may affect their oligomerisation, DNA-binding properties and interactions with other proteins, adding further level of complexity to understanding functions of these versatile proteins (Dilweg & Dame, 2018).


A
P
ON
P
P
fimS
fimB
fimE
fimA
IRL
IRR
IHF
IHF
H-NS
LRP
P
OFF
P
Rho
fimS
fimB
fimE
fimA
IRL
IRR
P

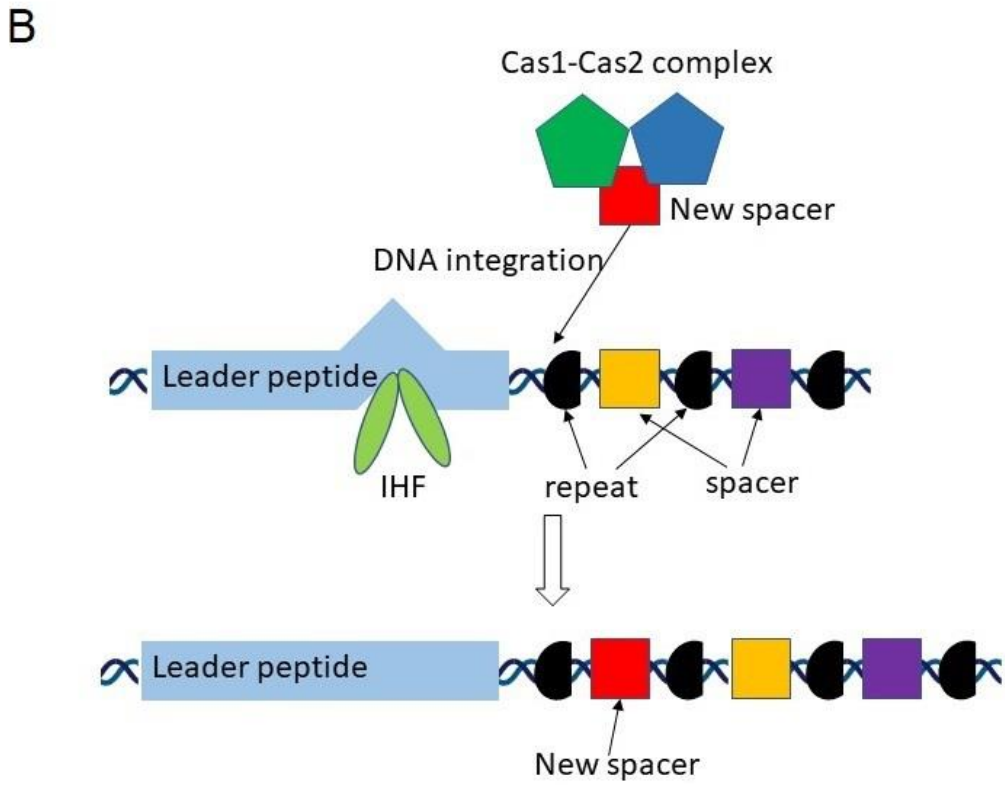

B
Cas1-Cas2 complex
New spacer
DNA integration
Leader peptide
IHF
repeat
spacer
Leader peptide
New spacer

Figure 1.4 Role of IHF in *fim* switch of *E. coli* and in CRISPR-*Cas*. a) The *FimS* invertible genetic switch of *E. coli*. To produce type 1 fimbriae and express it on its surface, *E. coli* needs to transcribe its *fim* operon located downstream of the *fimS* switch. The promoter that drives transcription of this operon in the ON state is located within the switch, flanked by the right and the left invertible repeats (IRL and IRR). Expression of the FimB and FimE tyrosine integrases in the ON state drives the recombination of the invertible repeats by the dominant FimE integrase, inverting the switch from the ON-to-OFF state. *Fim* operon is not transcribed in the OFF state, due to its promoter facing the other direction. Rho-dependent terminator stops transcription of the *fimE* prematurely, allowing FimB integrase to drive recombination of the switch from the OFF-to-ON state. Binding of IHF and LRP at their cognate sites allows the inversion in both directions by enabling response to supercoiling. H-NS is inhibitory to *fimB* and *fimE* transcription in the temperature-sensitive way. Green dashed arrows indicate sites of action of the proteins. b) CRISPR array of the CRISPR-*Cas* system. IHF binding at the leader peptide causes bending that promotes specific recognition of the leader peptide by the Cas1-Cas2 integrase and subsequent spacer acquisition. Based on (Bogue & Dorman, 2016; Nuñez *et al.*, 2016).

Table 1.1 Nucleoid-associated proteins of *E. coli* and *Salmonella*.

Name	DNA bending	DNA bridging	Self-aggregation	Binding specificity	Subunit composition	Reference
IHF	Yes	Yes	-	(A/T)ATCAANNNNTT(A/G)	Heterodimer	(Swinger & Rice, 2004; Yoshua *et al.*, 2020)
HU	Yes	-	-	Non-specific with preference for AT-rich or curved DNA	Heterodimer or homodimer	(Guo & Adhya, 2007)
H-NS	-	Yes	-	AT-rich DNA and TCGATAAATT	Homodimer or heterodimer with StpA	(Dame *et al.*, 2006)
StpA	-	Yes	-	AT-rich DNA	Homodimer or heterodimer with H-NS	(Zhang *et al.*, 1996)
Fis	Yes	Yes	-	AT-rich flanked by GC pairs	Homodimer	(Cho *et al.*, 2008)
Dps	-	-	Yes	DNA secondary structures	Dodecamer	(Almirón *et al.*, 1992)
MukB	-	Yes	-	Preference for ssDNA	Homodimer	(Niki *et al.*, 1991)
LRP	Yes	Yes	-	(T/C)AG(A/T/C)A(A/T)ATT(A/T)T(A/T/G)CT(A/G)	Monomer or octamer or hexadimer	(de los Rios & Perona, 2007)

Name	DNA bending	DNA bridging	Self-aggregation	Binding specificity	Subunit composition	Reference
CbpA	-	-	Yes	AT-rich and curved DNA	Homodimer or heterodimer with CbpM	(Cosgriff *et al.*, 2010)
CbpB	-	-	-	Curved DNA	Monomer	(Kakeda *et al.*, 1995)

1.1.4 DNA gyrase and other topoisomerases

Apart from NAPs, DNA superhelicity is the other major organising compaction factor of the nucleoid. The overall underwound nature of a chromosome, stated in section 1.1.2, means that to achieve a minimal energy state, a chromosome adopts a writhed conformation that absorbs the energy excess: this is referred to as negative supercoiling (Higgins & Vologodskii, 2015). This transformation of DNA shape that works in combination with the action of NAPs results in the overall compaction of the nucleoid (Dorman & Dorman, 2016). The opposite state when a chromosome is overwound is referred to as positive supercoiling. Not all organisms have their genome negatively supercoiled, some archaea that dwell in hyperthermal environments keep their genomes positively supercoiled as an adaptation to the heat stress (Lipscomb *et al.*, 2017).

The ability of DNA supercoils to store energy that can be used in other DNA transactions, implies that the supercoiling state of a chromosome is tightly linked to the energy state of a cell. Indeed, the intracellular ATP/ADP ratio determines the supercoiling state of DNA (van Workum *et al.*, 1996). This is because supercoiling is influenced by the action of enzymes called topoisomerases that are powered by ATP hydrolysis. There are four topoisomerases in *E. coli* and *Salmonella* (Table 1.2). It is a combinatory action of all topoisomerases that sets DNA supercoiling at both global and local levels. Two of these, DNA gyrase and Topoisomerase IV (Topo IV), are ATP-dependent type-II topoisomerases, while Topoisomerases I and III are ATP-independent type-I topoisomerases. The type of topoisomerase is based on the mechanism that the enzyme uses to change the linking number of the DNA duplex – in two steps and creating double-stranded DNA break (type II), or in one step and introducing single-stranded break in DNA (type I) (Champoux, 2001).

DNA gyrase of bacteria is composed of GyrA an GyrB subunits that are arranged in a heterotetrameric configuration. It belongs to the type IIa topoisomerase family sharing homology to Topo IV and Topoisomerase II of

Table 1.2 DNA topoisomerases of *E. coli* and *Salmonella*.

Name	Gene (s)	M_r (kDa)	Function (s)	References
Type I				
Topo I	*topA*	97	Relaxation of negative supercoils (ATP-independent).	(Bates & Maxwell, 2007)
Topo III	*topB*	73.2	Decatenation, weak DNA relaxation (ATP-independent).	(Perez-Cheeks *et al.*, 2012)
Type II				
DNA gyrase	*gyrA*/ *gyrB*	105/ 95	Introduction of negative supercoils, removal of positive supercoils (both ATP-dependent), DNA relaxation (ATP-independent).	(Higgins *et al.*, 1978)
Topo IV	*parC*/ *parE*	75/ 70	Decatenation of daughter chromosomes, removal of positive and negative supercoils (ATP-dependent).	(Bates & Maxwell, 2007)

yeast and higher eukaryotes and has a unique function among other topoisomerases to introduce negative DNA supercoils that makes it indispensable (Champoux, 2001). Despite a detailed understanding of gyrase's structure and mechanism, a crystal structure of a full DNA gyrase heterotetramer of *E. coli*, complexed with DNA, was obtained only recently (Vanden Broeck *et al.*, 2019a). GyrB contains an ATP-binding site at its N-terminus and a C-terminus that contains a DNA capture domain and is responsible for the interaction between the subunits and a TOPRIM fold that allows Mg^{2+} metal ion binding. The N-terminus half of GyrA contains the active site tyrosine within so-called CAP domains, while the C terminal half contains GyrA-box motif within a β-pinwheel structure that is essential for DNA wrapping (Champoux, 2001; Vanden Broeck *et al.*, 2019a) (Fig.1.5a).

The process of introducing a negative supercoil by DNA gyrase is precisely orchestrated (Fig. 1.5b). A DNA molecule can be thought of as being composed of a portion in which dsDNA break is made (gated or G-segment) and a portion that is passed through this break in order to introduce a supercoil (transport or T-segment). The dynamic heterotetramer can provide a high-affinity binding site for the G-segment DNA as well as three gated cavities in order to guide the T-segment: N-gate (opening between the N-termini of ATPase domain of GyrB subunits), DNA-gate (opening where the T segment passes through the G-segment) and the C-gate (opening between GyrA subunits that provides exit for the T-segment). The series of gate opening-closing events that happen at the expense of two molecules of ATP are shown on Fig. 1.5b. The 4 bp staggered dsDNA cleavage of the G-segment is accompanied by the formation of a transient phosphodiester bond between the 5' of the cleaved DNA and the active site tyrosines (Champoux, 2001; Vanden Broeck *et al.*, 2019a).

Apart from introducing negative supercoils into a relaxed DNA substrate, DNA gyrase can use the same mechanism to remove positive supercoils at a rate that exceeds that of the negative supercoil introduction (Ashley *et al.*, 2017). This is important when positive supercoiling is introduced in front of the moving DNA or RNA polymerases during replication and transcription. Resolving supercoiling patterns that are generated by the moving replication fork during

chromosomal replication and transcription apparatus during transcription is an important function carried out by topoisomerases. While DNA gyrase removes positive supercoils generated in front of the moving polymerases, Topo I and topo IV are responsible for the relaxation of negative supercoils that are generated behind the moving RNA polymerase (Fig. 1.6a) (Liu & Wang, 1987; Chong *et al.*, 2014). Topo IV is responsible for decatenation of newly replicated DNA behind the moving DNA polymerase (Fig. 1.6b) (Koster *et al.*, 2010; Kraemer *et al.*, 2019). Malfunction of any of the components results in stalling of the RNA or DNA polymerase and subsequent interruption of the corresponding reaction.

The mechanism of action of Topo IV is identical to that of DNA gyrase, except the substrate is not a single DNA molecule, but different molecules. While DNA gyrase is essential, due to its unique function, Topo IV can be absent from some bacteria, for example most *Mycobacteria*, *Campylobacter jejuni*, *Deonococcus radiodurans*. Decatenation in these bacteria is likely to be carried out by DNA gyrase (Manjunatha *et al.*, 2002). In bacteria with linear genomes, such as *Streptomyces*, Topo IV is present, although it is not required for decatenation of linear chromosomes. In these bacteria Topo IV is required for decatenation of spontaneously-forming circular replicons that form due to telomere-telomere interactions and circular plasmids (Huang *et al.*, 2013). DNA gyrase is estimated to be present at 600 copies per *E. coli* cell of which about 300 molecules are DNA-bound at any given time, presumably catalysing supercoiling adjustment reactions (Stracy *et al.*, 2019). Of these only about 12 enzymes are associated with each replisome and ensure efficient resolving of positive supercoils, introduced by very fast-moving replication machinery (travelling at about 1000 base pairs per second) (Reyes-Lamothe *et al.*, 2008), by performing multiple rounds of catalysis without dissociating from DNA substrate (Stracy *et al.*, 2019). In line with gyrase's activity in removal of positive supercoils in front of RNA polymerase, high association of DNA gyrase binding sites with highly transcribed operons was found in a recent ChIP study (Sutormin *et al.*, 2019). Two other topoisomerases of enteric bacteria are monomeric, ATP-independent Topo I and Topo III. Both enzymes make a transient cut in a single strand of DNA and use the energy stored in DNA to

drive the passage of the intact strand through the cut (Champoux, 2001). While the main function of Topo I is resolving negative supercoils (Fig 1.6a), Topo III is mainly involved in decatenation but has a weak DNA relaxation activity too (Perez-Cheeks *et al.*, 2012). DNA gyrase can also relax DNA independently of ATP, but this activity is evident only in the absence of ATP (Williams & Maxwell, 1999).

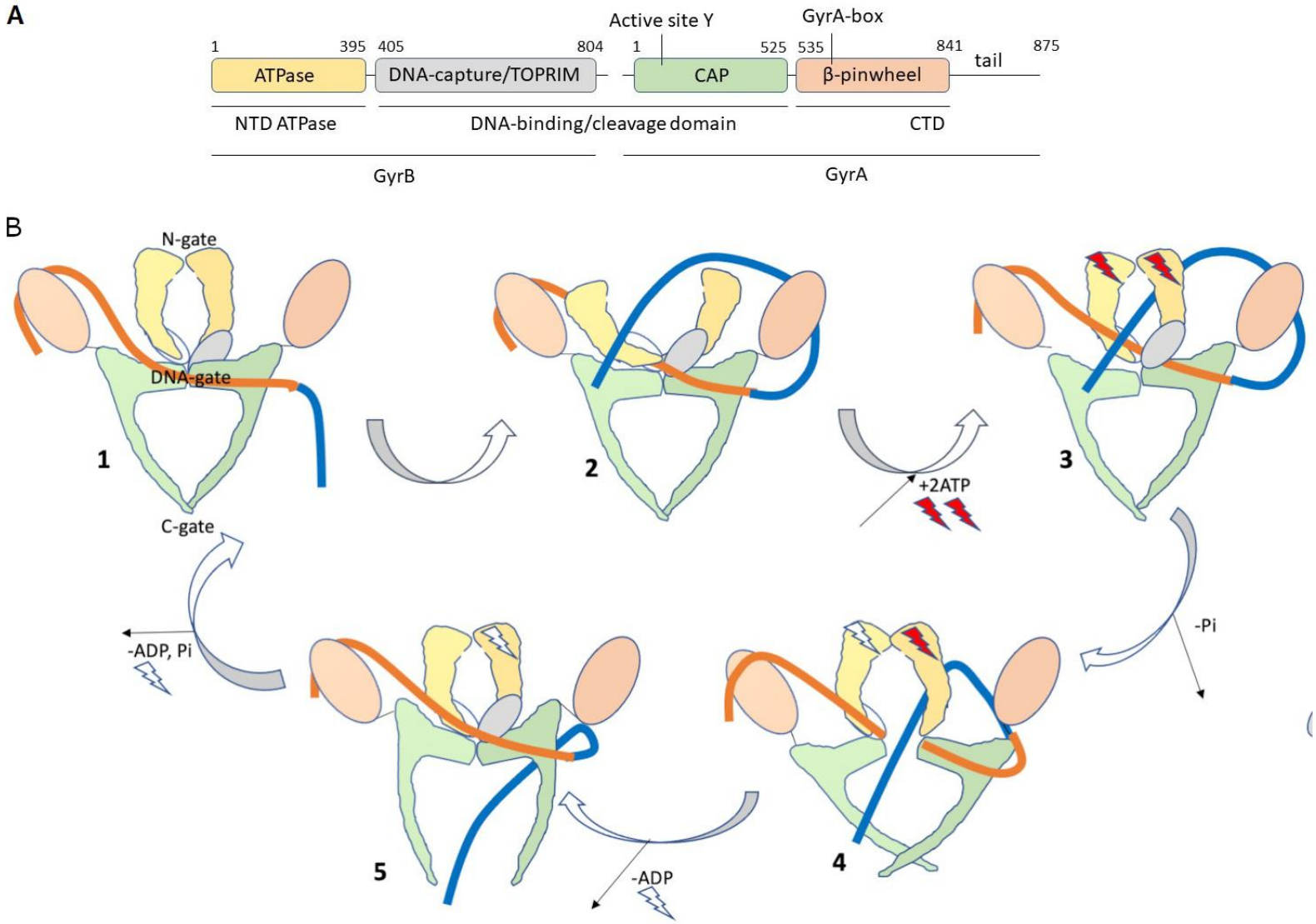

A
Active site Y
GyrA-box
1
395
405
804
1
525
535
841
875
ATPase
DNA-capture/TOPRIM
CAP
β-pinwheel
tail
NTD ATPase
DNA-binding/cleavage domain
CTD
GyrB
GyrA
B
N-gate
DNA-gate
1
2
3
+2ATP
C-gate
-ADP, Pi
-Pi
5
4
-ADP

Figure 1.5 DNA gyrase structure and function. a) Schematic representation of the major domains of the GyrA-GyrB complex. Not to scale. b) The process of introduction of a negative supercoil by DNA gyrase. Colours of the subunits correspond to those in the panel a. G-segment and T-segment DNA are shown in orange and blue, respectively. (1) All gates are closed and G-segment DNA binds to the DNA gyrase. (2) The N-gate opens, and the T-segment DNA passes through it guided by wrapping around the CTD of GyrA. (3) The N-gate closes and two ATP molecules bind to ATPases of each GyrB subunit. (4) The DNA gate opens, and the G-segment is cleaved with aid of the active site tyrosines at the expense of 1 ATP molecule that is converted to ADP+P_i. (5) The DNA gate closes, the G-segment is repaired the C-gate opens through which the T-segment exits the enzyme at the expense of the other ATP molecule. Based on (Champoux, 2001).

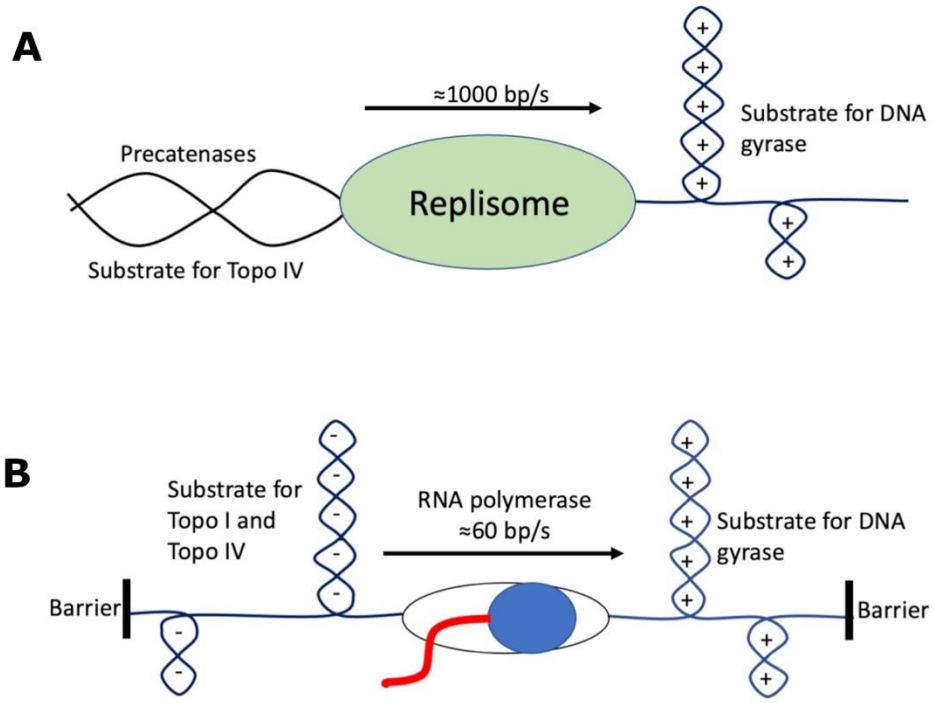


Figure 1.6 The topological environment of DNA replication and transcription. a) Movement of RNA polymerase during transcription in a piece of DNA restricted by barriers to supercoil diffusion introduces positive supercoils ahead and negative supercoils behind. DNA gyrase acts to resolve positive supercoils, while Topo I and Topo IV act in conjunction to resolve negative supercoils. b) Movement of replisome creates a region of positive supercoiling ahead that is resolved by DNA gyrase, and precatenases of the newly replicated DNA behind that are resolved by the action of Topo IV. Based on (Dorman, 2020).

1.1.5 *Salmonella* infection

Salmonella is an important pathogen that diverged from *E. coli* about 100 million years ago (Doolittle *et al.*, 1996) and is distinguished from it mainly by a number of horizontally acquired genetic features whose expression is the basis of *Salmonella's* pathogenicity. The species *Salmonella enterica* account for most human infections. The nomenclature of *Salmonella* is very complex – *Salmonella enterica* is further categorised into over 2000 Serovars that are defined by somatic (O) and flagellar (H) antigens (Brenner *et al.*, 2000). Human-restricted Serovars Typhi and Paratyphi cause typhoid fever in humans – condition characterised by systemic infection that was an important cause of death until the invention of antibiotics (Parry *et al.*, 2002). Not-typhoidal serovars such as Typhimurium and Enteritidis are responsible for nontyphoidal salmonellosis (NTS), both non-invasive and invasive (iNTS). NTS is characterised by an acute gastroenteritis with substantial liquid loss, while iNTS is a febrile systemic disease (Kurtz *et al.*, 2017). All of these diseases are associated with poor hygiene and contaminated water sources. They pose a threat in the developing world but can also cause outbreaks in the developed world and are associated with natural disasters and humanitarian crises (Kurtz *et al.*, 2017; Gilchrist & MacLennan, 2019).

There are at least 17 horizontally-acquired regions in *Salmonella* called *Salmonella* Pathogenicity Islands (SPIs) that are responsible for different stages of virulence (Ilyas *et al.*, 2017). Some encode type III secretion systems (T3SS), pathogenicity effectors and their regulators that mediate the bacterium's invasion and intracellular survival (SPI-1 and SPI-2). Others mediate adhesion to epithelial cells (SPI-4) or immune evasion (SPI-12) or take part in multiple virulence-related mechanisms (SPI-5) (Ilyas *et al.*, 2017). There is an interplay between different SPIs. For example, effectors of SPI-2 are encoded on SPI-5, on SPI-12 or on prophages. Their regulation is highly complex and is mediated by regulators encoded within SPIs and in the core genome at both transcriptional and post-transcriptional levels (Kröger *et al.*, 2012; Martínez *et al.*, 2014). Being horizontally acquired, SPIs have a high AT-content. This implies that they are silenced by H-NS to prevent unnecessary

and costly expression and expressed only when environmental conditions are permissive to do so (Stoebel *et al.*, 2008).

Once *Salmonella* cells are ingested by a host, they have to withstand the acidic environment in the stomach and competition from commensal bacteria in the gut, before reaching a dedicated entry point at the lumen of small intestine. This means that to establish a successful infection, large numbers of *Salmonella* cells must be ingested (Kurtz *et al.*, 2017). After *Salmonella* colonise the gut lumen, they promote inflammation in order to survive. This is a strange notion, given that a rapid innate immune response of a host is effective at killing pathogens. However, *Salmonella* evolved to evade the immune response and thrive with its aid (Kurtz *et al.*, 2017; Peterson & Artis, 2014). Induction of SPI-1 expression in the gut lumen results in active invasion of intestinal epithelial cells via the expression of T3SS-1 needles that inject SPI-1 effectors into a host cell, remodelling its actin cytoskeleton and inducing surface ruffling and subsequent uptake of *Salmonella* (Galán, 1996). After multiple bacteria gain access to the epithelial cell, they induce the normal state of the epithelial cell's surface to avoid an excessive immune response that could be detrimental for *Salmonella*. A SPI-1 independent mode of entry is also possible and occurs via Microfold cells (M cells) – a type of cell that performs sampling and uptake of luminal antigens and bacteria for the normal activity of the immune system or via routine luminal sampling by dendritic cells (Fig. 1.7) (Peterson & Artis, 2014).

The immediate innate immune response to *Salmonella* encompasses powerful cytokine secretion, neutrophil influx, and recruitment of other immune cells, such as macrophages and dendritic cells. In part, this aids clearing the infection, however, it also induces widespread tissue damage and associated fluid loss (Broz *et al.*, 2012). Following emergence from the basolateral membrane of epithelial cells, *Salmonella* encounter macrophages that are recruited in large numbers to the site of inflammation. Macrophages are phagocytic cells that kill engulfed pathogens by a combination of acidic pH, reactive oxygen and nitrogen species, hydrogen peroxide and lysozyme. However, followed by phagocytosis or SPI-1-induced active entry of *Salmonella* and its localisation inside a *Salmonella*-containing vacuole (SCV), these mechanisms fail to eradicate the pathogen. The acidic pH that bacteria

encounter inside a SCV is an environmental cue for expression of SPI-2. H-NS repression of SPI-2 is lifted by the combined action of SPI-2-encoded and core genome regulators and results in a number of measures that prevent macrophage-mediated bacteria killing mechanisms. SPI-2 expression prevents lysosomal fusion with the SCV, arrests oxidative burst and nitric oxide production (Figueira & Holden, 2012). *Salmonella's* own cytosol acidifies to pH 5-6 (Chakraborty *et al.*, 2015). Negative supercoiling during adaptation to acid stress is inhibited (Colgan *et al.*, 2018), resulting in the relaxation of *Salmonella's* genome (O'Croinin *et al.*, 2006). This results in de-repression of SPI-2 with the aid of enhanced OmpR binding (O'Croinin *et al.*, 2006; Quinn *et al.*, 2014). The main product of SPI-2 expression is the T3SS-2 needle that ejects the relevant effectors into the cytosol of the macrophage, mediating maturation of SCV to achieve *Salmonella* replication and eventual release (Figueira & Holden, 2012). After bursting of a macrophage and emergence, *Salmonella* cells are available to proceed with the macrophage infection cycle again or disseminate to the bloodstream, causing infection of organs and bacteremia in a permissive host (Fig. 1.7) (Ilyas *et al.*, 2017).

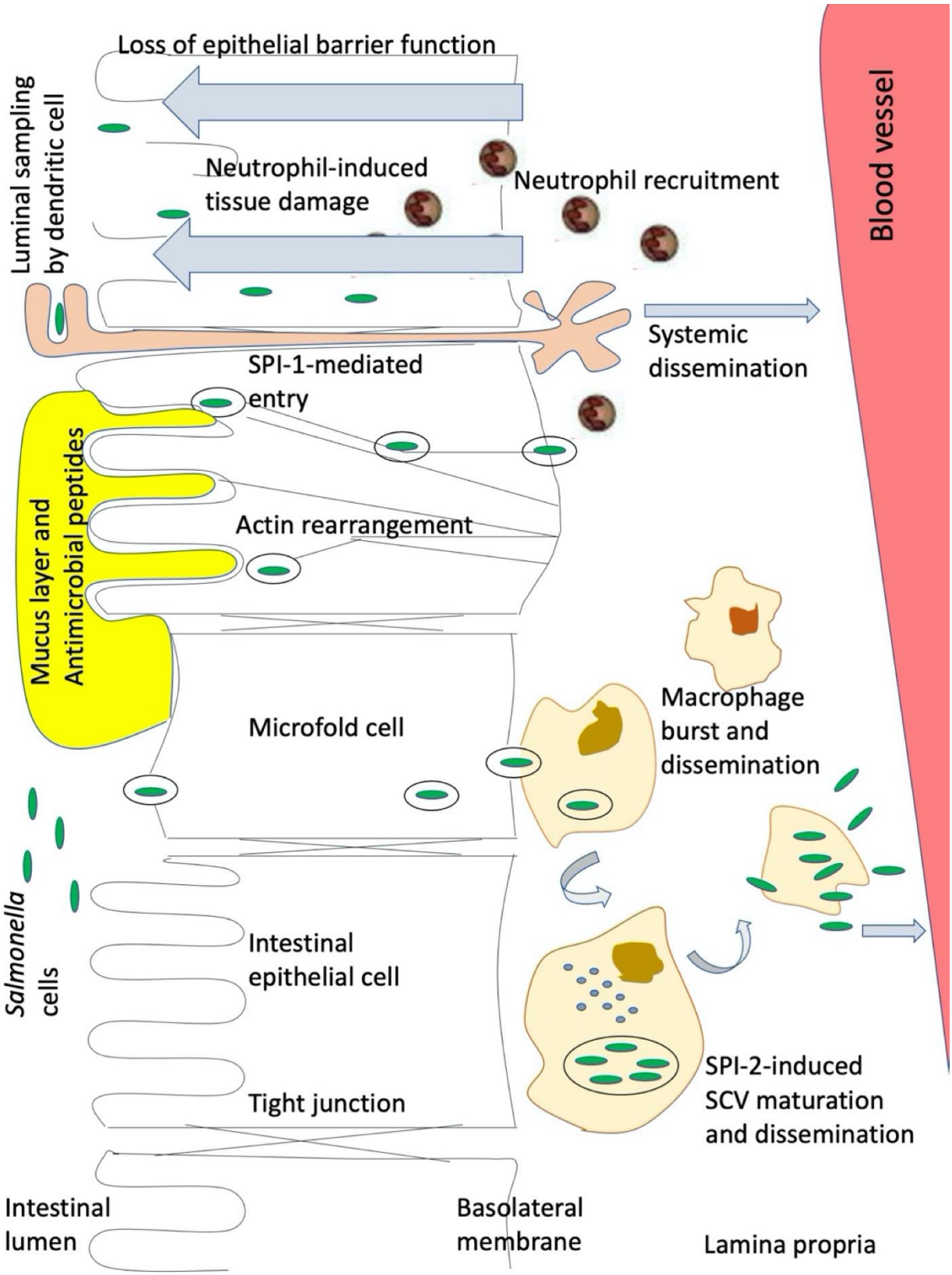

Loss of epithelial barrier function
Luminal sampling by dendritic cell
Neutrophil-induced tissue damage
Neutrophil recruitment
Blood vessel
Systemic dissemination
SPI-1-mediated entry
Mucus layer and Antimicrobial peptides
Actin rearrangement
Macrophage burst and dissemination
Microfold cell
Salmonella cells
Intestinal epithelial cell
SPI-2-induced SCV maturation and dissemination
Tight junction
Intestinal lumen
Basolateral membrane
Lamina propria

Figure 1.7 *Salmonella* infection. *Salmonella* can enter lamina propria from intestinal lumen via SPI-1-mediated active entry that is accompanied by the actin rearrangement of epithelial cell, via M cells or via dendritic cells. In lamina propria *Salmonella* is phagocytosed by macrophages or enters them actively and strives inside a *Salmonella*-containing vacuole (SCV). Expression of SPI-2 mediates *Salmonella* replication, SCV maturation and, eventually, macrophage burst. *Salmonella* can disseminate locally or systemically. Neutrophil recruitment results in killing of *Salmonella*, as well as tissue damage and loss of epithelial cell barrier function.

1.1.6 Chromosomal gene positioning.

An elaborate spatiotemporal organisation of the nucleoid implies that there may be meaningful conserved patterns in the way genes are arranged along a chromosome. Indeed, the degree of conservation of the chromosomal gene order is very high among gammaproteobacteria – the bacterial class to which both *E. coli* and *Salmonella* belong (Sobetzko *et al.*, 2012). While the degree of conservation between the right and the left replicores is very low, in agreement with the previous study (Alokam *et al.*, 2002), the relative position on the *oriC*-Ter axis is highly conserved. Moreover, the order of binding sites of various global regulators, such as sigma factors of RNA polymerase, DNA gyrase and major NAPs, displayed patterns that can be described as gradients along the *oriC*-Ter axis (Sobetzko *et al.*, 2012). All of it suggests that the *oriC*-Ter axis is an important internal reference for bacteria.

The importance of the *oriC*-Ter coordinates comes from the nature of DNA replication. In exponentially growing bacteria, there is a higher copy number of genes located near *oriC*. This is because a new round of replication is initiated before the previous round is completed (Schmid & Roth, 1987). This gives an opportunity for the origin-proximal genes to be expressed at a higher level during rapid growth. Therefore, multiple genes whose expression is required for optimal operation during rapid growth are located near the origin. For example, RpoD – a housekeeping RNA polymerase sigma factor, is exploited during rapid growth and *rpoD* is located near the *oriC*. Similarly, rapid growth sets a requirement for the high levels of translation, therefore, *rrn* operons that include multiple genes encoding rRNA, that can comprise up to 80% of cellular RNA during rapid growth, are located near the *oriC* (Gyorfy *et al.*, 2015). A recent genome-wide study of constitutive promoters of *E. coli* determined that those promoters that are highly dependent on the growth rate are preferably localized near *oriC* to exploit the benefits of a higher gene copy number (Yubero & Poyatos, 2020). There is also a related rough correspondence between the temporal expression of some proteins and their chromosomal position. For example, FIS – a NAP that is required for regulation of expression of multiple genes during the lag and early exponential phases of growth is encoded near *oriC*, while Dps – a NAP that is involved in the protection of DNA against

damage during the late stationary phase of growth is encoded near the terminus. Due to the limited ability of molecules to diffuse in the viscous cytoplasm, translation is coupled to transcription and protein products tend to have their targets located in the relative vicinity (McGary & Nudler, 2013; Schavemaker *et al.*, 2018; Montero Llopis *et al.*, 2010). For example, housekeeping and stress-related sigma factors RpoD and RpoS are encoded in the NS-Left closer to *oriC* and closer to the Ter, respectively. Their genetic targets are accordingly associated with *oriC* and Ter-proximal regions on a chromosome (Sobetzko *et al.*, 2012; Cameron *et al.*, 2017). The precise orchestration of the macrodomain movement inside a cell during the cell cycle ensures that proteins can reach their targets efficiently (Dame *et al.*, 2011).

Another important parameter to consider is gene orientation. During chromosomal replication one DNA strand is synthesised continuously (the leading strand), while the other strand is synthesised discontinuously via ligation of Okazaki fragments (the lagging strand). There is a large bias towards overrepresentation of genes in the leading strand, thus, co-oriented with the movement of replication fork. Since both transcription and replication operate on the same DNA template, this is inevitable that RNA polymerase stands on a way of a progressing replication fork (Liu & Alberts, 1995). The outcome of the collision between the polymerases depends on the direction of RNA polymerase. While co-oriented collisions do not lead to serious consequences, head-on collisions may slow the progression of the replication fork (Mirkin & Mirkin, 2005). Therefore, it is not surprising that genes in the lagging strand are underrepresented. It is, however, surprising that there are still genes located on the lagging strand, many of them are highly transcribed. In fact, essential genes but not highly expressed ones are biased to the leading strand. The lack of enrichment of highly expressed genes on a leading strand is likely to be due to the fact that it is the collision between the replication fork and RNA polymerase at the promoter of a gene per se that can slow down the replication fork. Therefore, number of RNA polymerases actively transcribing a gene or an extent to which a gene is expressed does not matter. The prevalence of the essential genes in the leading strand is because such an arrangement decreases the chance of formation of truncated peptides and gene expression

stochastic noise that would lead to deleterious effects when involving genes that encode proteins with essential functions (Rocha, 2008).

Genes on a chromosome can be located separately, as individual transcription units, each with its own promoter and flanking regulatory sites. Alternatively, they can be arranged in co-oriented group called an operon that shares a common promoter and all regulatory sites. The advantages of the operonic arrangement have been a subject for debates since *lac* operon was discovered in *E. coli* (Jacob & Monod, 1961). Multiple theories of operon development in prokaryotes have been formulated. Some of them have not been supported with sufficient evidence (Lawrence, 2003). For example, the natal theory suggests that operons arose by gene duplication and can explain the origin of eukaryotic gene clusters, such as *hox* genes, but not that of prokaryotic operons (Lawrence & Roth, 1996). The repair theory suggests that when genes, whose protein products contribute to one function are located in *cis*, mutations that accumulate can be repaired in one round of homologous recombination due to physical proximity (Hartman *et al.*, 1960). However, it does not explain how the operon assembled in the first place. According to the Fisher model, operons are products of frequent recombination that brings together genes with similar functions; once clustered, the selective pressure is high for the clusters to persist (Lawrence, 1997). However, bacterial chromosomes are remarkably stable, not taking into account horizontal gene transfer (HGT) and deletions. Very small amounts of genetic change have been introduced into the genome of *E. coli* by SNPs as compared to HGT (that is a main source of genetic change in bacteria) since it diverged from *Salmonella* (Lawrence & Ochman, 1997). Chromosomal inversions are also very unlikely to happen (unless the inversion is over the Ter region) since they are surprisingly more deleterious than insertions of foreign DNA or deletions (Rocha, 2008).

The most popular theories of operon formation are the regulatory theory and the selfish operon theory. However, these also have advantages and disadvantages when considered independently. The regulatory theory postulates that operons exist because of advantages of co-regulation (Osbourn & Field, 2009). The arguments in its favour include the higher selection imposed on genes within an operon, the advantage of encoding proteins that

are involved in the same function near each other due to the transcription-translation coupling in prokaryotes and the ease of controlling stoichiometries of such proteins (Rocha, 2008). This explains well the selective pressures behind the maintenance of the operons, but does not explain how they arose (Lawrence, 1997). Another caveat is that the regulatory theory predicts that operons should be composed of genes that encode proteins involved in the same function or part of the same complex. In reality, many operons code for unrelated proteins and optimal expression levels of genes in an operon could be different (Price *et al.*, 2006). The selfish operon model is conceptually different in that it suggests that operons form due to selection acting not on organisms but on genes (Lawrence & Roth, 1996). In order to increase their fitness, operons are transferred between prokaryotes horizontally. HGT does not only explain the origin of operons, but also that of other multi operon clusters, such as SPIs. One of the predictions of this model is that operons are likely to be composed of genes involved in functions that are useful for the recipient but not essential, such as utilisation of additional carbon source, or biosynthesis of a biomolecule that is novel for an organism, or virulence (Lawrence, 1997). This is because essential genes cannot be lost or gained easily via HGT. However, essential genes can frequently be found as part of operons, and this cannot be explained by the selfish operons model. Therefore, both models are likely to partially contribute to the creation and the maintenance of operons in bacteria.

The ability of a gene to express in sufficient levels and at the right time is a crucial consideration with regard to the gene positioning. Local superhelicity of DNA is one of the factors that affects gene expression, and orientation of the neighbouring genes is an important parameter that controls it (Sobetzko, 2016). For example, for a gene whose promoter is activated by relaxation of DNA supercoiling, a position between convergent genes would be optimal, since positive supercoils accumulate in front of the RNA polymerase during transcription (Chong *et al.*, 2014). The chromosome is not uniformly supercoiled (Lal *et al.*, 2016). This is due to activity of DNA gyrase, the binding sites of which are not distributed evenly throughout the chromosome, but associated with the *oriC*, replication forks and highly transcribed regions (Sobetzko *et al.*,

2012; Sutormin *et al.*, 2019). The chromosomal gene positions, especially those that are supercoiling-sensitive, are expected to be collinear with the distribution of DNA gyrase. Since NAPs such as FIS, IHF, HU and others are able to constrain local DNA supercoiling, gene positions also have to be considered in the context of neighbouring binding sites of these NAPs.

In addition to the importance of the gene dosage effect on chromosomal gene expression due to the *oriC* proximity (Soler-Bistué *et al.*, 2017; Slager & Veening, 2016), chromosomal location can also affect gene expression in ways that are independent of gene dosage, exemplified by the silencing of promoters inserted at the transcriptionally silent Extended Protein Occupancy Domains (tsEPOD) that overlap with NAP-bound regions (Bryant *et al.*, 2014). Binding of a gene promoter by global regulators can also affect its expression independently of the *oriC* proximity, for example, transcription silencing effect of H-NS binding at the AT-rich regions of the genome (Brambilla & Sclavi, 2015).

1.1.7 Synthetic biology

The exciting field of synthetic biology is at the interface of engineering and molecular biology. It is still in its infancy and requires more knowledge than we have already about gene regulatory networks and the organisation of DNA material in single-celled organisms as it is based on manipulations with DNA within living organisms. In order to construct artificial chromosomes or to modify an existing one to yield an organism with enhanced useful properties, we should be fully aware of the rules in place that shaped the structure of existing chromosomes, so that an outcome of placing a certain gene in certain locus would be fully predictable. Although an extensive knowledge exists on how different parts of organisms operate, the understanding of the overall regulatory principles and design is incomplete (Dorman, 2020). Studies of subjects described in the earlier sections of this chapter help connecting the dots of the overall picture.

There are several approaches that are used in synthetic biology. All of them, to some extent, involve taking the existing biological modules, rearranging them, and creating artificial genetic circuits, thereby constructing artificial genetic units or entire organisms (Bashor *et al.*, 2010). The modules in question are the genes, operons and horizontal transfer elements. The resemblance of biological modules to some existing electrical systems inspired creation of biological oscillators that mimic circadian rhythms (Tigges *et al.*, 2009), cell-base amplified biosensors (Hicks *et al.*, 2020), or short artificial gene networks with predicted behaviour (Ellis *et al.*, 2009). However, attempts to create more complex system often fail due to insufficient knowledge about hidden regulatory sites or introduction of unfavourable DNA secondary structures upon artificial alignment of the modules (Bartley *et al.*, 2017).

Another approach is the construction of a fully synthetic organism. Viral genomes have been constructed first, such as the small RNA genome of poliovirus. Chemical synthesis of viral cDNA led to effective *de novo* synthesis of infectious poliovirus (Cello *et al.*, 2002). In just a few years, synthesis of the prokaryotic genome had become possible. The strategy used was synthesising a genome of one organism and transferring it to the other organism. Bacteria of

genus *Mycoplasma* have been useful for this application due to their small genome (*Mycoplasma genitalium* has the smallest genome of all culturable bacteria) and defined number of essential genes (Glass *et al.*, 2006). The chromosome of *M. genitalium* was used as a frame to synthesise the first synthetic prokaryote by the complete transfer of the artificial genome to yeast *Saccharomyces cerevisiae* (Gibson *et al.*, 2008). Interestingly, when a minimal genome of another *Mycoplasma* – *M. mycoides* was synthesised, it was found that out of the 473 essential genes, 149 had no known biological function. This highlights that even when constructing a minimalistic synthetic organism under the tightly controlled laboratory conditions, the functionality of almost a third of the genomic content remains obscure (Hutchison *et al.*, 2016). It is likely that the first synthetic eukaryote will be *Saccharomyces cerevisiae*. The Synthetic Yeast Genome project is now underway and aims to construct all 16 chromosomes of *S. cerevisiae* (Pretorius & Boeke, 2018). It will answer fundamental questions related to genome organisation and properties of chromosomes. The eventual synthetic yeast might play an important practical role in wide range of industrial fermentations and could be precisely tailored for specific applications.

Synthetic biology has endless applications and employs different biological systems. The natural properties of bacteriophage; such as the ability to package the genetic information that can be tailored, simplicity of genome, host-specificity and high diversity; make phages a highly useful tool for synthetic biology applications (Lemire *et al.*, 2018). Advances in DNA assembly, genetic circuit wiring and CRISPR/*Cas* enable engineering of entire microbial communities for selected biotechnology applications (McCarty & Ledesma-Amaro, 2019). With number of articles found on the PubMed and dedicated to synthetic biology rising steeply over the past decade, this field is guaranteed to yield many more major advancements in the nearest future.

1.1.8 Aims of this work.

Having introduced the phenomenon of gene position conservation in bacteria, the importance of nucleoid structure, DNA supercoiling and associated gene regulation, I would like to highlight the major aims of this work, prior to introducing them in detail in the following chapters.

There are two major studies in the focus of this thesis seeking to alter chromosomal positions of the genes that encode important global regulators: IHF and DNA gyrase. These have been inspired by other recent studies that altered the chromosomal positions of the promoters, ORFs and the entire genes that encoded various bacterial global regulators (Brambilla & Sclavi, 2015; Gerganova *et al.*, 2015; Fitzgerald *et al.*, 2015; Bogue *et al.*, 2020). None of these studies found essentiality of existing gene positions, however, the repositionings affected multiple properties of bacteria due to the central regulatory roles of the proteins that those genes encoded. Other previous studies that looked at the regulon and global binding of IHF and DNA gyrase were useful sources of information when interpreting the results of the current work (Mangan *et al.*, 2006; Prieto *et al.*, 2012; Sobetzko *et al.*, 2012; Sutormin *et al.*, 2019). The IHF and the DNA gyrase studies were developed side-by-side and, therefore, share methodology. For example, strain construction employed the same method and some phenotypic tests were identical, however, there were other experiments that tested alterations associated with a change in phenotypes characteristic of IHF or DNA gyrase and, therefore, unique to each of the two studies. Testing the significance of the chromosomal gene positions of the genes that encode IHF and DNA gyrase might suggest the optimal positions for these genes on a chromosome of a synthetic bacterium and determine the drawbacks of the alternative configurations.

To summarise, the aims of the studies described in this thesis are:

- To construct *Salmonella* strains with the reciprocal ORF exchanges of the naturally separate genes *ihfA* and *ihfB* that encode the NAP IHF, as well as strains with artificial *ihfAB* and *ihfBA* operons.
- To construct *Salmonella* strains with the artificial DNA gyrase operons *gyrAB* or *gyrBA*.

- To study the phenotypic effects of the chromosomal repositionings by examining phenotypes associated with IHF and DNA gyrase.
- To study the patterns of expression of the repositioned genes at both mRNA and protein levels.
- To study the effect of the gene repositionings on global gene expression at the level of transcription and/or protein production.
- To gain more insight into the biology of IHF and DNA gyrase and the reasons behind the separate positioning of the genes that encode each of these proteins.

Chapter 2:

Materials and Methods

2.1 Chemicals and growth media

2.1.1 Chemicals

All chemicals were obtained from Sigma-Aldrich (Arklow, Ireland) unless otherwise stated. Where required media and chemicals were sterilized by autoclaving (121°C for 20 minutes). If not suitable for autoclaving, chemical solutions were filter-sterilized using 0.22 µm pore-sized filter (Merck, Millipore). Water used throughout was sterile distilled water. Nuclease-free water was used for DNA work, DEPC-treated water (Fisher, Geel, Belgium) was used for RNA work.

2.1.2 Antibiotics

Antibiotics were typically prepared in 1000x stock concentrations and used at the following fixed working concentrations: carbenicillin (ampicillin) – 100 µg/ml in water, kanamycin – 50 µg/ml in water, chloramphenicol – 35 µg/ml in ethanol, rifampicin – 500 µg/ml in methanol. Antibiotics that were used at variable working concentrations were prepared at the following stock concentrations: novobiocin – 25 mg/ml in water, ciprofloxacin – 10 mg/ml in 0.1N HCl, nalidixic acid – 50 mg/ml in water, coumermycin (Santa Cruz Biotechnology, Dallas, USA) – 5 mg/ml in DMSO. All antibiotic solutions were filter-sterilized using 0.22 µm filter and stored in aliquots at -20°C, except coumermycin that was kept at 4°C.

2.1.3 Growth media

Bacterial strains were routinely grown in sterile lysogeny broth (LB) Miller formulation: (1% tryptone, 0.5% yeast extract (both supplied by BD, Le Pont de Claix, France), 1% NaCl (Acros Organics, Geel, Belgium) supplemented with antibiotics where required. 1.5% agar (BD, Le Pont de Claix, France) was used for pouring plates and 0.3% agar was used for motility measurements. Green agar was made from 1.5% agar base with the addition of 0.0625% (w/v) alizarin yellow, 0.006% aniline blue and 0.84% glucose.

N-minimal medium was used in experiments where the effect of magnesium ion concentrations was assessed. Buffer N was made first (Nelson & Kennedy, 1971): 5 mM KCl, 7.5 mM $(NH_4)_2SO_4$ (BDH, Poole, England), 0.5 mM K_2SO_4, 1 mM KH_2PO_4 (Fisher, Geel, Belgium), and 0.1 M pH 7.4 Tris-HCl (Acros Organics, Geel, Belgium). To make full N-minimal medium, buffer N was supplemented with 0.2% glucose (BDH, Leuven, Belgium), 1 µg/ml thiamine, 0.5 mM L-histidine and 10 mM or 10 µM $MgCl_2$ (Acros Organics, Geel, Belgium) to obtain high or low Mg^{2+} concentration respectively.

2.2 Bacterial and eukaryotic strains and culture conditions

2.2.1 Bacterial strains

All *Salmonella* strains were made in *Salmonella enterica* serovar Typhimurium SL1344 background and this strain was used as a wild type (WT). Plasmids were sourced from a carrier *Escherichia coli* DH5α strain. All bacterial strains used are listed in Table 2.1. The major strains are drawn schematically (Fig. 2.1).

2.2.2 Bacterial culture condition

Bacterial strains were stored as 35% glycerol stocks at -80°C and freshly streaked on agar plates for each biological replicate. A single colony was inoculated into 4 ml LB broth and grown for 18 h. This overnight culture was subcultured into fresh 25 ml LB broth normalizing to an OD_{600} of 0.003, unless otherwise stated, and grown to the required growth phase. The standard growth conditions for all bacterial strains were 37°C, 200 rpm, unless otherwise stated.

For culturing in minimal medium, overnight cultures were prepared as described above. 1 ml of overnight culture was washed three times with N-minimal medium of the required Mg^{2+} concentration to remove nutrients, subcultured into minimal medium of the corresponding Mg^{2+} concentration in a total volume of 25 ml and grown for 24 h to pre-condition the bacteria. The pre-conditioned culture was subcultured into 25 ml of fresh minimal medium normalized to an

Table 2.1 Bacterial strains used in this study.

Strain name	Genotype/Description	Reference
SL1344 (WT)	*rpsL hisG*	(Hoiseth & Stocker, 1981)
ihfA::*cat*	Chloramphenicol resistance cassette inserted downstream of the *ihfA* ORF	This study
ihfB::kan	Kanamycin resistance cassette inserted downstream of the *ihfB* ORF	This study
2x*ihfA*	2 *ihfA* copies, *ihfB* ORF replaced with the *ihfA* ORF that has the kanamycin resistance cassette inserted downstream – Δ*ihfB.*	This study
2x*ihfB*	2 *ihfB* copies, *ihfA* ORF replaced with the *ihfB* ORF that has the chloramphenicol resistance cassette inserted downstream – Δ*ihfA*	This study
OrfSwap$^{ihfA\text{-}ihfB}$	*IhfA* and *ihfB* ORFs reciprocally exchanged	This study
Δ*ihfA*	*IhfA*::kan/ *ihfA* ORF deleted, except the start codon and the last 21 bp, by inserting the kanamycin resistance cassette	This study
Δ*ihfB*	*IhfB*::kan/ *ihfB* ORF deleted, except the start codon and the last 21 bp, by inserting the kanamycin resistance cassette	This study
ihfAB	*IhfAB* operon under the control of the *ihfA* promoter	This study
ihfBA	*IhfBA* operon under the control of the *ihfB* promoter	This study
gyrA::kan	Kanamycin resistance cassette inserted downstream of the *gyrA* ORF	This study

gyrB::kan	Kanamycin resistance cassette inserted downstream of the *gyrB* ORF	This study
gyrBA	*GyrBA* operon under the control of the *gyrB* promoter	This study
*gyrBA**	*GyrBA* operon under the control of the *gyrB* promoter but with SNPs	This study
gyrAB_gyrB	*GyrAB* operon under the control of the *gyrA* promoter, two copies of *gyrB*	This study
SL1344_*ihfA*::3x-FLAG	Insertion of 3 copies of the FLAG epitope tag downstream of *ihfA*, deleting its stop codon, in the SL1344 background	This study
SL1344_*ihfA*::8x-*myc*	Insertion of 8 copies of the Myc epitope tag downstream of *ihfA*, deleting its stop codon, in the SL1344 background	This study
SL1344_*ihfB*::3x-FLAG	Insertion of 3 copies of the FLAG epitope tag downstream of *ihfB*, deleting its stop codon, in the SL1344 background	This study
SL1344_*ihfB*::8x-*myc*	Insertion of 8 copies of the Myc epitope tag downstream of *ihfB*, deleting its stop codon, in the SL1344 background	This study
OrfSwap$^{ihfA\text{-}ihfB}$ _*ihfA*::3x-FLAG	Insertion of 3 copies of the FLAG epitope tag downstream of *ihfA*, deleting its stop codon, in the OrfSwap$^{ihfA\text{-}ihfB}$ background	This study
OrfSwap$^{ihfA\text{-}ihfB}$ _*ihfA*::8x-*myc*	Insertion of 8 copies of the Myc epitope tag downstream of *ihfA*, deleting its stop codon, in the OrfSwap$^{ihfA\text{-}ihfB}$ background	This study
OrfSwap$^{ihfA\text{-}ihfB}$ _*ihfB*::3x-FLAG	Insertion of 3 copies of the FLAG epitope tag downstream of *ihfB*,	This study

	deleting its stop codon, in the OrfSwap*ihfA-ihfB* background	
OrfSwap*ihfA-ihfB* _*ihfB*::8x-*myc*	Insertion of 8 copies of the Myc epitope tag downstream of *ihfB*, deleting its stop codon, in the OrfSwap*ihfA-ihfB* background	This study
SL1344_*ihfA*::3x-FLAG *ihfB*::8x-*myc*	Insertion of 3 copies of the FLAG epitope tag downstream of *ihfA*, deleting its stop codon, and with insertion of 8 copies of the Myc epitope tag downstream of *ihfB*, deleting its stop codon, in the SL1344 background	This study
SL1344_*ihfA*::8x-*myc* *ihfB*::3x-FLAG	Insertion of 8 copies of the Myc epitope tag downstream of *ihfA*, deleting its stop codon, and with insertion of 3 copies of the FLAG epitope tag downstream of *ihfB*, deleting its stop codon, in the SL1344 background	This study
OrfSwap*ihfA-ihfB* _*ihfA*::8x-*myc*_*ihfB*::3x-FLAG	Insertion of 8 copies of the Myc epitope tag downstream of *ihfA*, deleting its stop codon, and with insertion of 3 copies of the FLAG epitope tag downstream of *ihfB,* deleting its stop codon, in the OrfSwap*ihfA-ihfB* background	This study
OrfSwap*ihfA-ihfB* _*ihfA*::3x-FLAG_*ihfB*::8x-*myc*	Insertion of 3 copies of the FLAG epitope tag downstream of *ihfA*, deleting its stop codon, and with insertion of 8 copies of the Myc epitope tag downstream of *ihfB*, deleting its stop codon, in the OrfSwap*ihfA-ihfB* background	This study

SL1344_P*prgH*::GFP	P*prgH-gfp*[LVA]/R::*cat*/ fusion of a *gfp* gene encoding a destabilised version of GFP to the SPI-1 promoter, P*prgH*	(Ibarra *et al.*, 2010)
OrfSwap$^{ihfA\text{-}ihfB}$ _P*prgH*::GFP	Fusion of a *gfp* derivative encoding a destabilised version of GFP to the SPI-1 promoter, P*prgH* in the OrfSwap$^{ihfA\text{-}ihfB}$ background	This study
Δ*ihfA*_P*prgH*::GFP	Fusion of a *gfp* derivative encoding a destabilised version of GFP to the SPI-1 promoter, P*prgH* in the 2x*ihfA* background	This study
Δ*ihfB*_P*prgH*::GFP	Fusion of a *gfp* derivative encoding a destabilised version of GFP to the SPI-1 promoter, P*prgH* in the 2x*ihfB* background	This study
*gyrBA*_P*prgH*::GFP	Fusion of a *gfp* derivative encoding a destabilised version of GFP to the SPI-1 promoter, P*prgH* in the *gyrBA* background	This study
SL1344_P*ssaG*::GFP	P*prgH-gfp*[LVA]/R::*cat*/ fusion of a *gfp* derivative encoding a destabilised version of GFP to the SPI-2 promoter, P*ssaG*	(Ibarra *et al.*, 2010)
OrfSwap$^{ihfA\text{-}ihfB}$ _P*ssaG*::GFP	Fusion of a *gfp* derivative encoding a destabilised version of GFP to the SPI-2 promoter, P*ssaG* in the OrfSwap$^{ihfA\text{-}ihfB}$ background	This study
Δ*ihfB*_P*ssaG*::GFP	Fusion of a *gfp* derivative encoding a destabilised version of GFP to the SPI-2 promoter, P*ssaG* in the Δ*ihfB* background	This study
Δ*ihfA*_P*ssaG*::GFP	Fusion of a *gfp* derivative encoding a destabilised version of GFP to the SPI-	This study

	2 promoter, P*ssaG* in the Δ*ihfA* background	
*gyrBA*_P*ssaG*::GFP	Fusion of a *gfp* derivative encoding a destabilised version of GFP to the SPI-2 promoter, P*ssaG* in the *gyrBA* background	This study
SL1344_*SL1483*::*cat*	Insertion of a chloramphenicol resistance cassette into the pseudogene *SL1483*	
OrfSwap$^{ihfA\text{-}ihfB}$ _*SL1483*::*cat*	Insertion of a chloramphenicol resistance cassette into the pseudogene *SL1483* in the OrfSwap$^{ihfA\text{-}ihfB}$ background	This study
ihfAB_SL1483::*cat*	Insertion of a chloramphenicol resistance cassette into the pseudogene *SL1483* in the *ihfAB* background	This study
ihfBA_SL1483::*cat*	Insertion of a chloramphenicol resistance cassette into the pseudogene *SL1483* in the *ihfBA* background	This study
gyrBA_SL1483::*cat*	Insertion of a chloramphenicol resistance cassette into the pseudogene *SL1483* in the *gyrBA* background	This study
*gyrBA*_SL1483*::*cat*	Insertion of a chloramphenicol resistance cassette into the pseudogene *SL1483* in the *gyrBA** background	This study

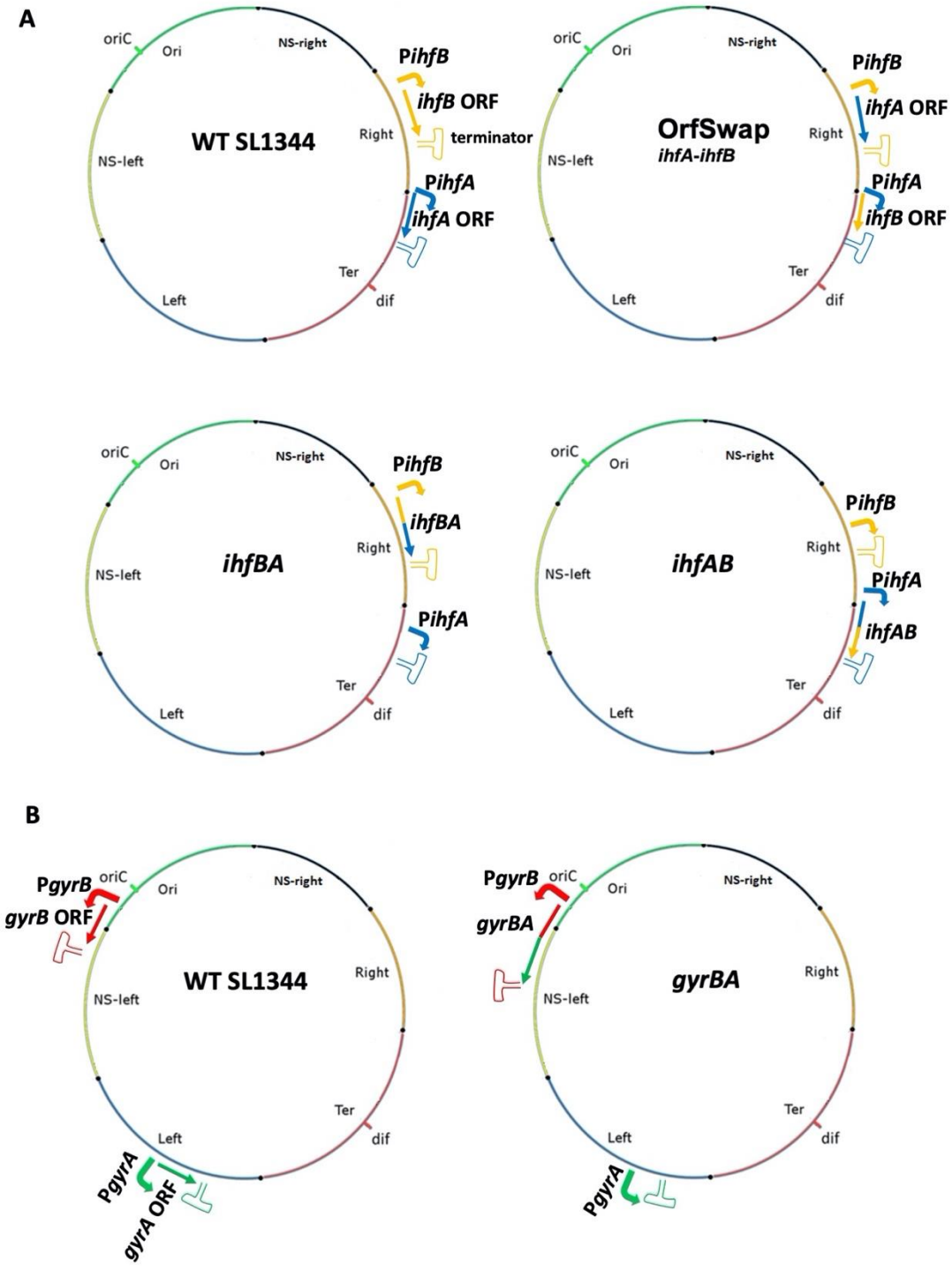


Figure 2.1 Major strains used in this work. a) The WT SL1344 and the strains with the repositionings of *ihfA* and *ihfB*. b) The WT SL1344 and the *gyrBA* strains. Positions of *oriC*, dif and macrodomains are shown. Promoter, ORF and terminator of the genes of interest are shown and colour-coded. Schematic drawings are not to scale.

OD_{600} of 0.03 and grown for a further 24 h to obtain a culture in the stationary phase of growth.

2.2.3 Mammalian cell lines

Murine macrophage RAW264.7 cells was obtained from Frederick J. Sheedy, School of Biochemistry & Immunology, Trinity College Dublin.

2.2.4 Mammalian cell culture conditions

RAW264.7 murine macrophages were maintained in Dulbecco's Modified Eagle's Medium (DMEM), (Sigma, catalogue number D6429) supplemented with 10% fetal bovine serum (FBS) in a humidified 37°C, 5% CO_2 tissue-culture incubator grown in 75 cm^3 tissue-culture flasks. When approximately 80% confluent growth was achieved, cells were split to a fresh flask. Cells within the 9-16 passage number range were used for infections. All media and PBS used for cell culture were pre-warmed to 37°C.

To split cells, old DMEM was removed and the monolayer was rinsed with 10 ml of sterile PBS. 10 ml of fresh DMEM was pipetted into the flask and the monolayer was scraped gently with a cell scraper to dislodge the cells. Scraped cells were centrifuged at 450 x g for 5 min on Eppendorf 5810R centrifuge, the cell pellet was resuspended in 5 ml DMEM+FBS. 1 ml of the cell suspension was added to 14 ml of fresh DMEM+FBS in a 75 cm^3 flask, gently rocked to mix and incubated at 37°C, 5% CO_2.

To seed cells for infection, cells were treated as for splitting. After resuspension in 5 ml DMEM+FBS, viable cells were counted on a haemocytometer using trypan blue exclusion dye. A 24-well tissue culture plate was filled with 500 µl DMEM+FBS. 1.5×10^5 cells were added to each well, gently rocked to mix and incubated at 37°C, 5% CO_2 for 24 h.

2.3 Oligonucleotides, plasmids and bacteriophages

2.3.1 Oligonucleotides

All oligonucleotides (primers) used in this study were designed by the author and ordered from Integrated DNA Technologies (Leuven, Belgium). Primers were delivered in lyophilised form and resuspended in nuclease-free water to a final concentration of 100 pmol/µl and stored at -20°C. The sequences of all primers used are listed in Table 2.2.

2.3.2 Plasmids

All plasmids used in this study are listed in Table 2.3. The plasmids were selected for by using antibiotics as appropriate.

2.3.3 Bacteriophage

Bacteriophage P22 HT 105/1 *int*-201 was used for generalized transductions throughout the study (Schmieger, 1972). Lysates prepared were filter-sterilized and stored at 4°C in the dark.

2.4 Monitoring bacterial growth

2.4.1 Monitoring bacterial growth in rich media by spectrophotometry

To measure growth characteristics of a bacterial culture, an overnight culture was normalized to the optical density at 600 nm (OD_{600}) of 0.003 in 25 ml of fresh LB broth and grown at the standard conditions for 24 h. OD_{600} was measured every 1 h for the first 3 hours and then every 30 min until 8 hours, the last reading was taken at 24 h. Measurements were taken using the Thermo Scientific BioMate 3S spectrophotometer in plastic cuvettes using fresh LB broth for 1:10 dilutions of culture when its OD_{600} exceeded 1.0.

Table 2.2 Oligonucleotide primers used in this study

Name	5'-3' sequence
Confirmation Primers:	
SL_ihfA_check_Pfwd	CAACCGTACACTCGAAGAAGAG
SL_ihfA_check_Prev	GAACGTTTCGTCGCTGTTG
SL_ihfB_check_Pfrw	GCAATGGCTGAAGCATTCAAAG
SL_ihfB_check_Prev	GACCGTCGTTATCTTCATAGACAC
gyrA_check_Pf2	GACTAAGGTAGCGGTAAATG
gyrA_check_Prev2	GTAGATGACGAAGAACTCG
gyrB_check_Pf2	CAACGAATCCATTCCGATG
gyrB_check_Prev2	CTGATGAGCAGACTGTAAC
gyrA check del R	GCATTGTCTGGCTGCATTC
gyrB check del R	CTTTGTCAGCGCAATTAGC
gyrA_midcheck	CGATGGTGTCATACACTG
gyrB_midcheck	GAACGGTCATGATCACTTC
qPCR primers:	
SL_ihfA_qPCR_Pf	GATAAGCTTGGGCTTAGCA
SL_ihfA_qPCR_Pr	GAGAGTTTCACCTGCTCAC
SL_2ihfB_qPCR_Pf	CAATCTCACATTCCCGCTAAG
SL_2ihfB_qPCR_Prev	GCGGATTTCAATACGCTCG
SL_btuC_qPCR_Pf	GGCTACACTTCTGAGCTTATG
SL_btuC_qPCR_Pr	GGTTCAGCAAGTGGGTTT
SL_ycaI_qPCR_Pf	GTCAGCGTATGCGTAGTATG
SL_ycaI_qPCR_Pr	CAAAGCAAACATGCCAGAC
RT_gyrB_F	CTCGTTCAGCTCGGTAATCAG
SL_gyrB_R	ATGATTGGTCGTATGGAGCG
SL_gyrA_qPCR_Pf	CAGCGGTACCGTGAAGAAA
SL_gyrA_qPCR_Prev	AGCATGACTTCGTCAGAACC
RT_hemX_F	CGCCTGACGGTATGTTTCTT
RT_hemX_R	CCCAACCAGGACGTCTATTTAC
SL_ihfB_qPCR_Pf	CATATGGCCTCGACTCTTG
SL_ihfB_qPCR_Pr	TCCTTCCAGTTCCACTTTATC

Fis_qPCR_Pf	TGACGTACTGACCGTTTCTACCGT
Fis_qPCR_Prev	ACGTCCTGACCATTCAGTTGAGCA
OmpR_qPCR_Pf	CTCTCGCGATAAGCTGATGAA
OmpR_qPCR_Prev	CCGGATCTTCTTCCACCATAC
FNR_qPCR_Pf	GTCTGGCCGCGTTTATTTAC
FNR_qPCR_Prev	GACCGTTAAGCCCAGATAGTT

***ihfA* ORF insertion into *infB* locus**

infA.cmR.Pfwd	GAGCCGGGTCGAAAACGCTTCGCCCAAAGAAG AGTAATCAGTGTAGGCTGGAGCTGCTTC
infA.cmR.Prev	GCGTATCTGCCGCAATACACCCTGATGGATGTTATG CCTGCATATGAATATCCTCCTTAG
infB.int.Pfwd	ACGGCTGCAGCCAATTTGCCTTTAAGGAACCGGAGGA ATCATGGCGCTTACAAAAGCTG
infB.int.Prev	CGGTGCTTTTTTCGGGTTCAAGTTTTGCGTTAAAA CCCTGCATATGAATATCCTCCTTAG

***ihfB* ORF insertion into *ihfA* locus**

ihfB.kanR.Pfwd	AGAACTGCGCGATCGCGCCAATATTTACGGTTA AGTTTTAGTGTAGGCTGGAGCTGCTTC
ihfB.kanR.Prev	CAAAACTTGAACCCGAAAAAAGCACCGTCAGGG TGCTTTTCATATGAATATCCTCCTTAG
ihfA.int.Pfwd	AAAAGAGCGATTCCAGGCATCATTGAGGGATTG AACCTATGACCAAGTCAGAATTGATTG
ihfA.int.Prev	GCAATACACCCTGATGGATGTTATGCCTGGATC TGACATATGAATATCCTCCTTAGTTCC

ihfB*::kan insertion downstream of *ihfA

ihfA.int.ihfB::kan_Pf	GAGCCGGGTCGAAAACGCTTCGCCCAAAGAAG AGTAATCATTGCCTTTAAGGAACCGGAG
ihfA.int.ihfB::kan_Prev	CGTATCTGCCGCAATACACCCTGATGGATGTTA TGCCTGGCATATGAATATCCTCCTTAG

ihfA*::Cm insertion downstream of *ihfB

ihfB.int.ihfA::Cm_Pf	TAAAGAACTGCGCGATCGCGCCAATATTTACGG
	TTAAGTTCAGGCATCATTGAGGGATTG
ihfB.int.ihfA::Cm_Prev	AGCACCCTGACGGTGCTTTTTTCGGGTTCAAGT
	TTTGCGTCCTGCATATGAATATCCTCC

Deletion mutations – kan^r insertions

Kan_ihfA_Pf	AAAGAGCGATTCCAGGCATCATTGAGGGATTGA
	ACCTATGGTGTAGGCTGGAGCTGCTTC
Kan_ihfA_Prev	ATGTTATGCCTGGATCTGATTACTCTTCTTTGGG
	CGAAGCCATATGAATATCCTCCTTAG
Kan_ihfB_Pf	GCTGCAGCCAATTTGCCTTTAAGGAACCGGAGG
	AATCATGGTGTAGGCTGGAGCTGCTTC
Kan_ihfB_Prev	TCAAGTTTTGCGTTAAAACTTAACCGTAAATATT
	GGCGCGCATATGAATATCCTCCTTAG
Kan_gyrA_Pf	CCCTCGCACAGCAATAACATTACTCGTCAGCGT
	CATCCGCCATATGAATATCCTCCTTAG
Kan_gyrA_Prev	CTTTGAATCCGGGATACAGTAGAGGGATAGCGG
	TTAGATGGTGTAGGCTGGAGCTGCTTC
Kan_gyrB_Pf	GGCCGGGGATTAAGGCAGGTTAAATATCGATAT
	TCGCTGCCATATGAATATCCTCCTTAG
Kan_gyrB_Prev	ACGGATTAACCCAAGATTAAATGAGCGAGAAAC
	GTTGATGGTGTAGGCTGGAGCTGCTTC

gyrB*::kan insertion downstream of *gyrA

Kan ins gyrB F	TCATGATGCCCCGGCCAACCAGCGGTAGGCCG
	GGGATTAAGTGCCACCTGCATCGAT
Kan ins gyrB R	GGAGAACGCCCTGAAAGCAGCGAATATCGATAT
	TTAACCTCATATGAATATCCTCCTT
gyrA.int.gyrB::Kan_Pf	GGCAAAGAAAAAGGGCCGGATATCCGGCCCTC
	GCACAGCAGTGTAGGCTGGAGCTGCTTC
gyrA.int.gyrB::Kan_Prev	AAGCGATGACGACGTTGCGGATGACGCTGACG
	AGTAATGTTAAATGAGCGAGAAACGT

gyrA::kan insertion downstream of _gyrB_

Kan ins gyrA F	GGCAAAGAAAAAGGGCCGGATATCCGGCCCTC GCACAGCAGTGCCACCTGCATCGAT
Kan ins gyrA R	AAGCGATGACGACGTTGCGGATGACGCTGACG AGTAATGTCATATGAATATCCTCCTT
gyrB.int.gyrA::kan_Pf	TCATGATGCCCCGGCCAACCAGCGGTAGGCCG GGGATTAAGTGTAGGCTGGAGCTGCTTC
gyrB.int.gyrA::kan_Prev	GGAGAACGCCCTGAAAGCAGCGAATATCGATAT TTAACCTAGTAGAGGGATAGCGGTT

Epitope tagging

SL_ihfA_Pf-flag	GTTAAAGAGCCGGGTCGAAAACGCTTCGCCCAA AGAAGAGGACTACAAAGACCATGACGG
SL_ihfA_Prev-flag	TGCCGCAATACACCCTGATGGATGTTATGCCTG GATCTGACATATGAATATCCTCCTTAG
SL_ihfB_Pf-flag	ACCGGGTAAAGAACTGCGCGATCGCGCCAATAT TTACGGTGACTACAAAGACCATGACGG
SL_ihfB_Prev-flag	CTGACGGTGCTTTTTTCGGGTTCAAGTTTTGCGT TAAAACCATATGAATATCCTCCTTAG
SL_ihfA_Pf-myc	TTAAAGAGCCGGGTCGAAAACGCTTCGCCCAAA GAAGAGGTCGGATCCAGTCTTCGTGAT
SL_ihfA_Prev-myc	GCCGCAATACACCCTGATGGATGTTATGCCTGG ATCTGAAATTCCGGGGATCCGTCGACC
SL_ihfB_Pf-myc	CCGGGTAAAGAACTGCGCGATCGCGCCAATATT TACGGTGTCGGATCCAGTCTTCGTGAT
SL_ihfB_Prev-myc	TGACGGTGCTTTTTTCGGGTTCAAGTTTTGCGTT AAAACAATTCCGGGGATCCGTCGACC
SL_ihfA_Pf-HA	TTAAAGAGCCGGGTCGAAAACGCTTCGCCCAAA GAAGAGTATGATGTTCCTGATTATGCT
SL_ihfA_Prev-HA	GCCGCAATACACCCTGATGGATGTTATGCCTGG ATCTGACATATGAATATCCTCCTTAGT
SL_ihfB_Pf-HA	CCGGGTAAAGAACTGCGCGATCGCGCCAATATT TACGGTTATGATGTTCCTGATTATGCT

SL_ihfB_Prev-HA	TGACGGTGCTTTTTTCGGGTTCAAGTTTTGCGTTAAAACCATATGAATATCCTCCTTAGT

For strain construction primers – the black portion is an annealing end, the red portion is an overhanging end. All primers were designed in this study.

Table 2.3 Plasmids used in this study

Plasmid name	Description	Reference
pKD3	Amp^R ($Carb^R$), Cm^R	(Datsenko & Wanner, 2000)
pKD4	Amp^R ($Carb^R$), Kan^R	(Datsenko & Wanner, 2000)
pKD46	Amp^R ($Carb^R$), λ Red genes *γ, β, exo* under the control of an arabinose inducible promoter	(Datsenko & Wanner, 2000)
pCP20	Amp^R ($Carb^R$), Cm^R, FLP recombinase expressing, temperature sensitive replicon	(Cherepanov & Wackernagel, 1995)
pUC18	Amp^R ($Carb^R$),	(Yanisch-Perron *et al.*, 1985)
pSUB11	Amp^R ($Carb^R$), Kan^R	(Uzzau *et al.*, 2001)
pBOP508	Kan^R	(Cho *et al.*, 2006)

Abbreviations: Amp^R ($Carb^R$), ampicillin (carbenicillin) resistance; Cm^R, chloramphenicol resistance; Kan^R, kanamycin resistance.

2.4.2 Monitoring bacterial growth in minimal medium by spectrophotometry

To measure growth characteristics of a bacterial culture in minimal medium with altered Mg^{2+} concentration, an overnight bacterial culture was washed in minimal medium with corresponding concentration of Mg^{2+} and pre-conditioned for 24 h, as in 2.2.2. The pre-conditioned culture was normalized to an OD_{600} of 0.03 in 25 ml of fresh medium in two flasks and OD_{600} was measured every 1 h beginning from 2 h until 8 h. Separate cultures were set up similarly to measure OD_{600} every hour from 8 h until 15 h. In this way the number of times each flask was opened and sampled was minimized to yield reliable measurements.

2.4.3 Monitoring bacterial growth in rich medium by plate viability counting

Growth characteristics of a bacterial culture in LB broth were also measured by viable count. The culture was grown in the same way as in 2.16.1, and an aliquot of culture was taken at 2 h, 4 h, 6 h, 8 h and 24 h, serially diluted and spread-plated on LB agar plates to give between 30 and 300 colonies after overnight incubation at 37°C. The colony counts were converted into bacterial colony forming units per millilitre (cfu/ml).

2.5 Nucleic acid preparation

2.5.1 DNA isolation for routine uses

Bacterial chromosomal DNA, for uses such as a template for polymerase chain reaction (PCR) and providing a standard curve in the quantitative polymerase chain reaction (qPCR), was extracted from overnight cultures using the chloroform method. 400 µl of the overnight culture was harvested by centrifuging in a table-top Eppendorf 5415R centrifuge at 16000 x g for 1 min. The cell pellet was resuspended in 200 µl of sterile distilled water and boiled on a heating block at 100°C for 5 min. The cell suspension was vortexed briefly to ensure cell wall disruption and centrifuged at 16000 x g for 5 min to remove

cellular debris. The supernatant was carefully transferred into a clean Eppendorf tube, 1 volume of chloroform (Fisher, Loughborough, UK) was added and vortexed for 30 sec to ensure complete mixing. The mixture was then centrifuged at 16000 x g for 10 min at 4°C. About 75% of the upper aqueous DNA-containing layer was transferred into a clean tube. The DNA concentration was measured as described in 2.5.5, diluted to 100 ng/µl and stored at -20°C.

2.5.2 Plasmid DNA isolation

Plasmid DNA was isolated with the aid the of QIAprep Spin miniprep kit (QIAGEN, Hilden, Germany) according to manufacturer's guidelines. For routine uses, plasmid DNA was extracted from a 4 ml overnight culture. Briefly, the kit used alkaline lysis to disrupt the cells and denature the DNA (both chromosomal and plasmid). The RNase component in the resuspension buffer removed any RNA. Lysis was followed by a brief neutralization step that allowed small circular plasmid DNA to reform, but this step was not long enough to renature large chromosomal DNA. Cellular debris was pelleted by centrifugation and the cleared lysate was applied to a silica membrane column by centrifugation during which dsDNA plasmid molecules bound to the membrane and denatured chromosomal DNA was washed away. Plasmid DNA was desalted through a series of washes, eluted with nuclease-free water and its concentration was determined as in 2.5.5.

2.5.3 DNA isolation for whole genome sequencing

To obtain high-quality chromosomal DNA for whole genome sequencing, a basic phenol-chloroform method was used (Sambrook & Russell, 2006). 2ml of an overnight culture were centrifuged at 16000 x g for 1 min to harvest cells and the cell pellet was resuspended in 400 µl of TE buffer pH 8 (100 mM Tris-HCl pH 8.0, 10 mM EDTA pH 8.0 (BDH, Poole, England)). 1% SDS and 2 mg/ml proteinase K were added and incubated for 2 h at 37°C to complete lysis. DNA was isolated by the addition of 1 volume of phenol pH 8.0 : chloroform : isoamyl

alcohol (25 : 24 : 1) (AppliChem, Darmstadt, Germany), thorough mixing and centrifugation at 16000 x g for 15 min at 4°C in phase-lock tube. The upper aqueous layer containing DNA was collected and the phenol: chloroform extraction was repeated two more times. To remove contaminants and to precipitate DNA, sodium acetate pH 5.2 at 0.3 M and isopropanol at 60% of the final volume were added and kept for 1 h at -20°C. DNA was pelleted by centrifugation at 16000 x g for 15 min at 4°C. The DNA pellet was washed with 70% ethanol, dried at 37°C until translucent and resuspended in 100 µl TE pH 8.0. The sample was electrophoresed on agarose gel to check for degradation as in 2.6.5 and DNA concentration was determined as in 2.5.5.

To remove RNA contamination from DNA samples, 100 mg/ml RNase A was added and incubated for 30 min at 37°C. Phenol-chloroform extraction was performed as above. To precipitate DNA, 0.5 M of ammonium acetate (Merck, Darmstadt, Germany) and a half volume of isopropanol were added and incubated for 2 h at -20°C. DNA was pelleted by centrifugation at 16000 x g for 15 min at 4°C. The DNA pellet was washed twice with 70% ethanol, dried at 37°C until translucent and resuspended in 50 µl water. The sample was run on an agarose gel to check for degradation as in 2.6.5 and the DNA concentration was determined as in 2.5.5.

2.5.4 RNA isolation for quantitative PCR

RNA for measuring gene expression by qPCR was isolated using an acidic phenol-chloroform method. An overnight culture was subcultured into 25 ml of fresh LB broth normalising to an OD_{600} of 0.003. The bacterial culture was grown to the required timepoint and mixed with 40% volume of 5% acidic phenol (pH 4.3) in ethanol and placed on ice for at least 30 min to stop transcription. The cells were harvested by centrifugation at 3220 x g for 10 min at 4°C and resuspended in 700 µl of TE buffer pH 8 containing 0.5 mg/ml lysozyme. 1% SDS and 0.1 mg/ml proteinase K were added and incubated for 20 min at 40°C to complete lysis. 1/10 volume of 3 M sodium acetate was added to precipitate RNA, 1 volume of 1:1 solution of acidic phenol and chloroform was added, mixed well on a vortex mixer and centrifuged at 16000 x

g for 15 min at 4°C to extract RNA into aqueous phase. To precipitate RNA the aqueous layer was harvested, mixed with 1 volume of isopropanol and incubated at -20°C for 1 hour. RNA was harvested by centrifugation at 16000 x g for 15 min at 4°C. The RNA pellet was washed with 70% ethanol and dried at 37°C until translucent. The total RNA was dissolved in 50 µl DEPC-treated water and its concentration was determined as in 2.5.5.

2.5.5 Nucleic acid concentration determination

Concentration of DNA and RNA extracted was determined by measuring absorbance at 260 nm on a DeNovix DS-11 spectrophotometer (Wilmington, US). Shape of the absorbance curve was ensured to have a clear peak at 260 nm. Purity of samples was assessed by the ratio of A_{260}/A_{280} – a measure of protein and phenol contamination and A_{260}/A_{230} – a measure of contaminants such as EDTA, where both should be as close as possible to 2. Only high-quality samples were chosen for further work.

2.6 Polymerase chain reaction and work with its products

2.6.1 Polymerase chain reaction

The Polymerase Chain Reaction (PCR) was used for *in vitro* amplification of DNA fragments (Saiki *et al.*, 1985). The reaction is based on the activity of a thermostable DNA polymerase enzyme to amplify DNA that was denatured by heating and primed by sequence-specific oligonucleotide primers using provided dNTPs and in presence of Mg^{2+} cofactor. Nuclease-free water was used throughout to preserve DNA integrity. Eppendorf, Nexus MasterCycle eco/gradient (Hamburg, Germany) PCR cyclers were used.

2.6.2 *Taq* polymerase chain reaction

For uses such as strain validation, colony screening for a particular DNA insertion, or checking for DNA contamination, PCR by *Taq* DNA polymerase (New England Biolabs, Ipswich, US) was used. DNA template for PCR was

prepared as in 2.5.1 or colony PCR was performed by using a single bacterial colony diluted in 20 µl water as a template. Reactions were carried out in 10 µl total volume containing 1x Standard *Taq* Reaction Buffer, 200 µM dNTPs, 1 µM of each forward and reverse primer, 1 µl of DNA or colony template, 0.25 U *Taq* DNA polymerase and water. The PCR reactions were prepared on ice and transferred into a thermocycler. The cycling conditions for *Taq* were as follows:

1. Initial denaturation. 95°C, 30 sec (10 min if colony PCR)
2. Denaturation. 95°C, 15 sec
3. Primer annealing. 45-68°C, 30 sec
4. Template extension. 68°C, 1 min per 1 kb.
5. Steps 2-4 repeated 30 times
6. Final extension. 68°C, 5 min
7. Hold. 4°C

Primer annealing temperature was determined from its sequence. Pairs of primers were designed to have similar lengths and a similar GC-content. For all PCR types, primer annealing temperatures were determined using the melting temperature calculator from New England Biolabs (tmcalculator.neb.com). An annealing temperature gradient was used to find the optimal annealing temperature.

2.6.3 High-fidelity polymerase chain reaction

For strain construction when no errors during DNA amplification are tolerated, high fidelity Phusion® DNA polymerase (NEB, Ipswich, US) was used. This enzyme has 50x higher fidelity than *Taq* due to its 3'→ 5' exonuclease proofreading activity (Frey & Suppmann, 1995). Plasmid template, extracted as in 2.4.2, or a PCR product, purified as in 2.5.4, were used as a template. Reactions were carried out in a 20 µl total volume containing 1x Phusion HF buffer, 200 µM dNTPs, 0.5 µM of each forward and reverse primer, 0.4 U Phusion DNA polymerase, variable amount of DNA template and water. The PCR reactions were prepared on ice and transferred into a thermocycler. The cycling conditions for Phusion were as follows:

1. Initial denaturation. 98°C, 30 sec
2. Denaturation. 98°C, 10 sec
3. Primer annealing. 45-72°C, 30 sec
4. Template extension. 72°C, 30 sec per 1 kb
5. Steps 2-4 repeated 35 times
6. Final extension. 72°C, 5 min
7. Hold. 4°C

If required, additional steps taken to improve the specificity and yield of the reaction: temperature gradient to find an optimal primer annealing temperature, template and primer titration to find the optimal concentration of either, addition of 3% DMSO to prevent DNA secondary structure formation.

2.6.4 Purification of PCR products

For the purpose of transformation into competent cells during strain construction and for Sanger sequencing, PCR products were purified using QIAquick PCR purification kit (QUIAGEN, Hilden, Germany) according to manufacturer's guidelines. Briefly, PCR was upscaled to obtain 100 µl of product. 5 volumes of a high-salt buffer were added, and DNA was bound to a silica membrane assembly by centrifugation. Ethanol buffer was used for washing and DNA was eluted with 50 µl water. DNA concentration was determined as in 2.5.5.

2.6.5 DNA agarose gel electrophoresis

To assess the quality and size of PCR products, samples were run on 1.5% agarose gels made with TAE buffer (40 mM Tris, 20 mM acetic acid (Fisher, Loughborough, UK), 1 mM EDTA). DNA was mixed with 10x loading dye (60% glycerol (Fisher, Loughborough, UK), 0.1% bromophenol blue, 50 mM Tris-Cl pH 7.5) prior to loading. HyperLadder 1 kb (Bioline, London, UK) was used as a molecular size marker. The samples were electrophoresed through the gel for about 40 min at 90 V, the gel was stained in GelRed Nucleic Acid Gel Stain (Baotou, USA) for 30 min and viewed under UV light.

2.7 DNA sequencing

2.7.1 Sanger sequencing of PCR products

Sanger sequencing was used when it was required to find out if the strain construction was going correctly, providing the nucleotide base sequences of those chromosomal loci where genetic rearrangements were introduced. The region of interest was amplified using *Taq* PCR as in 2.6.2, the reaction was upscaled, the PCR product was purified as in 2.6.4 and sent off for sequencing along with product-specific primers. The sequencing was performed by Source BioScience (Tramore, Ireland). The resultant sequence reads were compared against the expected sequence using SnapGene software (GSL Biotech, San Diego, USA; available at snapgene.com).

2.7.2 Whole genome sequencing

Whole genome sequencing was performed on final versions of the constructed strains to ensure that no compensatory mutations were introduced into their genomes. The sequencing was performed by the Sanger Institute (Hinxton, Cambridge, UK) and MicrobesNG (Birmingham, UK) using Illumina next generation sequencing technology on samples prepared as in 2.5.3. The output reads were assembled using Velvet (Zerbino, 2010) and aligned to the reference SL1344 sequence NC_016810.1 Breseq software (Deatherage & Barrick, 2014).

2.8 Bacterial transformation with plasmid DNA

2.8.1 Preparation of electrocompetent cells

An overnight culture was diluted 1:100 in 25 ml of LB broth and grown to an OD_{600} of 0.4-0.6 at the standard growth conditions. Half of the culture was centrifuged at 3220 x g for 10 min at 4°C. The bacterial pellet was resuspended in 10 ml ice-cold water and centrifuged at the same conditions. The pellet was resuspended in 1 ml of ice-cold water and centrifuged at 16000 x g for 10 min

at 4°C. The wash was repeated twice and the pellet was resuspended in 200 µl of ice-cold water.

2.8.2 Electroporation of electrocompetent cells with plasmid DNA

40 µl of electro-competent cells were mixed with 200 ng, 400 ng or 600 ng of plasmid on ice and added to a pre-chilled 2 mm electrode gap width electroporation cuvette (Cell Projects, Kent, UK). The cuvette was wiped dry, placed into the Gene Pulser chamber (Bio-Rad, Dublin, Ireland) and an electric shock was applied at 2.5 kV. 1 ml of warm LB broth was added immediately, and the cells were allowed to recover for 1 h at the standard growth conditions. 150 µl aliquots were spread on LB agar selection plates containing an appropriate antibiotic. Negative controls that contained no DNA were included.

2.9 Lambda-red recombination

2.9.1 Preparation of linear DNA for transformation

Linear DNA for transformation was prepared by amplifying a region of DNA from a plasmid or a chromosome that is to be moved into a different location by PCR with a pair of primers. Each primer contained an approximately 20 b.p. section annealing to the 5**'** or the 3**'** end of the region to be moved and an had an overhanging end of 40 b.p. complementary to a region on *Salmonella* chromosome where the linear DNA was to be inserted. Phusion PCR was used as in 2.6.3, the reaction was scaled up and its products were purified as in 2.6.4.

2.9.2 Induction of *γ, β* and *exo* gene expression in pKD46

To transform *Salmonella* with linear DNA and ensure its incorporation into the genome λ-red-mediated homologous recombination was used as described previously (Datsenko & Wanner, 2000). The method depends on the activities of genes comprising the λ-red system encoded on the plasmid pKD46. This plasmid was transformed into a recipient strain, as in 2.8.2; the transformants

were recovered and grown at 30°C, as pKD46 has temperature sensitive replication. An overnight culture of the pKD46-carrying recipient strain was diluted 1:100 in 50 ml of LB broth and grown to an OD_{600} of 0.5 at 30°C, 200 rpm in the presence of carbenicillin and 100 mM arabinose to induce the expression of the λ-red system of pKD46.

2.9.3 Preparation of electrocompetent cells for transformation with linear DNA and electroporation

To achieve higher transformation efficiency required for the incorporation of linear DNA fragments into the *Salmonella* genome, a 50-ml recipient strain culture was grown to an OD_{600} of 0.5 as in 2.9.2. The cells were harvested by centrifugation at 3220 x g for 20 min at 4°C and subjected to a series of washes at 4°C with 20 min incubation on ice between the washes. The following washes were carried out: 50 ml of water, 25 ml of water and 2 ml of 10% glycerol. Finally, the cells were resuspended in 400 µl of 10% glycerol and electroporated as in 2.8.2. Cells remaining after spread-plating were kept at room temperature overnight.

2.9.4 Confirming the integration of linear DNA

Candidate colonies that appeared on selective agar plates after overnight incubation were transferred onto an appropriate selection plate with sterile toothpicks and incubated overnight at standard conditions to eliminate false positives. If no candidate colonies were found, the remaining cells that were kept at room temperature overnight were harvested by centrifugation, resuspended in fresh LB broth, spread plated onto selective plates and incubated overnight. Each toothpicked candidate colony was screened by PCR as in 2.6.2, the PCR products were electrophoresed and visualised as in 2.6.5. The successful integration of linear fragments was determined by correlating sizes of the PCR products obtained using the same pair of primers between candidate strains and the original unmodified strain with the predicted sizes. A

candidate colony that was shown to have a correct integration by PCR was sequenced by Sanger sequencing as in 2.7.1.

2.10 Generalised transduction with bacteriophage P22

2.10.1 P22 lysate preparation

To obtain a high-titre lysate of a strain containing a desired mutation, an overnight culture of a mutant strain was diluted 1:1000 in 25 ml of LB broth with suitable selection and grown to an OD_{600} of about 0.1. Next, 20 µl of P22 HT 105/1 *int-201* lysate prepared on SL1344 WT was added and the culture was incubated for a further 4 h at the standard growth conditions. 500 µl of chloroform was added and incubated for 10 min at room temperature to kill any surviving *Salmonella* cells. The lysate was centrifuged at 3220 x g for 10 min at 4°C to pellet cell debris. The lysate in the top layer was filter-sterilized with 0.22 µm pore size filter and stored at 4°C in the dark.

2.10.2 Transduction by bacteriophage P22

P22 transductions were used to transfer mutations into a clean genetic background (marker rescue) during strain construction and for the introduction of selectable mutations made in a *Salmonella* background into any other *Salmonella* recipient. 100 µl of an overnight culture of a recipient strain was mixed with 100 µl of neat, 1:10 and 1:100 P22 lysate dilutions prepared on a strain harbouring a donor fragment. Recipient-only and lysate-only controls were included. The mixtures were incubated statically for 1 h at 37°C to allow for transduction to proceed, plated on appropriate selection plates and incubated overnight. Colonies from the plate with highest dilution of lysate were picked for phage elimination.

2.10.3 Elimination of P22 pseudo-lysogens

To ensure that no phage is carried over into the transduced *Salmonella* culture, pseudo-lysogens must be distinguished from true lysogens. Each transductant

colony was streaked for single colonies on a green agar plate and incubated overnight at 37°C. A light-green colony that is free of pseudo-lysogens was picked for the second passage on a green plate, while dark blue colonies were avoided. Dark blue colour development was due to pH indicator changing colour in response to leakage of cellular contents caused by P22 lysis of the pseudo-lysogens.

2.11 Removal of antibiotic resistance cassettes using pCP20-mediated FLP site-specific recombination

During strain construction kanamycin and chloramphenicol resistance cassettes were used when selecting for insertion of linear DNA into chromosome. These antibiotic resistance cassettes had to be removed to enable further strain construction steps. This was done using pCP20-mediated FLP site-specific recombination (Cherepanov & Wackernagel, 1995). A strain from which a resistance cassette had to be removed was transformed with pCP20, as in 2.8.2, with recovery at 30°C. If removing a Kan^R cassette, 150 µL of cells were plated on a chloramphenicol or a carbenicillin plate. If removing a Cm^R cassette, 50 µL of cells were plated on carbenicillin plate. The plates were incubated at 30°C overnight. One colony was used to set up an overnight culture at 37°C with no antibiotics (passage 1). The passage 1 was streaked for single colonies on a plate without antibiotics, a carbenicillin plate to check for the removal of pCP20 and on a kanamycin plate, if removing a Kan^R, or on a chloramphenicol plate, if removing a Cm^R. The plates were incubated overnight at 37°C. If there was growth on any of the selection plates, a colony from the plate with no antibiotics was used to set up an overnight culture at 30°C without antibiotics (passage 2). The same plating strategy as above was used to check for pCP20 and resistance cassette removal after the second passage. When there was no growth on any of the selection plates, the strain was stored as a glycerol stock at -80°C.

2.12 Manipulation of RNA *in vitro*

2.12.1 DNase treatment of RNA

DNase I treatment was used to remove DNA contamination from RNA samples extracted as in 2.5.4. RNA was diluted to 20 µg in 80 µl, denatured at 65°C for 5 min and kept on ice. 1x DNase I buffer including $MgCl_2$ and 10 U DNase I (ThermoFisher Scientific, Waltham, US) were added and incubated for 45 min at 37°C. 100 µl of 1:1 acidic phenol : chloroform was added to DNase I digestion samples, mixed and transferred to a phase-lock tube. RNA was extracted by centrifugation at 16000 x g for 12 min at 15°C. The upper aqueous layer was harvested and RNA was precipitated by adding 2.5 volumes of 30:1 ethanol : 3 M sodium acetate pH 6.5 for 2 h or overnight at -20°C. RNA was harvested by centrifugation at 16000 x g for 30 min at 4°C. The RNA pellet was washed with 70% ethanol and dried at 37°C until translucent. The total RNA was dissolved in 30 µl DEPC-treated water and its concentration was determined as in 2.5.5. RNA was checked for DNA contamination by the end point PCR.

2.12.2 Determining RNA integrity by gel electrophoresis

RNA was tested for integrity before and after DNase treatment by electrophoretic separation on a Hepes/triethanolamine (HT) agarose gel (Mansour & Pestov, 2013). 2% agarose gel was made in 1x HT Buffer (30 mM Hepes (Fluorochem, Hadfield, UK), 30 mM triethanolamine) with addition of 400 mM of formaldehyde. 1x HT buffer was used as a running buffer. 5 parts of RNA was mixed with 10 parts of HT loading dye (30 mM Hepes, 30 mM triethanolamine, 0.5 mM EDTA, 0.02% bromophenol blue, 400 mM formaldehyde, 10% glycerol) and 5 parts of formamide. The samples were denatured at 65°C for 5 min, cooled down on ice for 5 min and loaded on a gel. The gel was run for 2 h at 60 V, stained in GelRed for 30 min and viewed under UV. Large rRNA species were assessed for integrity.

2.12.3 cDNA synthesis

Total extracted and DNase I treated RNA was converted to cDNA using GoScript™ Reverse Transcription System kit (Promega) according to manufacturer's guidelines. Briefly, 400 ng of DNase treated RNA was mixed with 1 µl of random primers in 5 µl reaction. The RNA-primer mix was denatured for 5 min at 70°C and cooled down on ice. A reverse transcription mix was assembled on ice and added to the RNA-primer mix. No-reverse-transcriptase and no-template controls were included. Primer annealing was carried out for 5 min at 25°C followed by the extension for 60 min at 42°C. 100 µl of nuclease-free water was added to obtain cDNA concentration of 3.33 ng/µl.

2.13 Assays based on quantitative polymerase chain reaction (qPCR)

2.13.1 cDNA quantification using Real Time qPCR

To determine the amount of mRNA transcript from a specific gene at a particular growth stage, total RNA was extracted as in 2.5.4, DNase treated as in 2.12.1 and converted into cDNA as in 2.12.3. Then, 5.33 ng of cDNA in 20 µl reaction was used as a template for Real Time quantitative PCR (RT-qPCR) using 1x FastStart Universal SYBR Green Master (ROX) (Roche, Mannheim, Germany) and gene-specific pair of primers (0.3 µM each) in 96-well plate format. The FastStart Taq DNA Polymerase contained in the Universal Master mix was used for PCR amplification of cDNA, while SYBR Green I dye emits fluorescence at 530 nm when it intercalates with PCR products. During PCR, the increase in SYBR Green I fluorescence is directly proportional to the amount of double stranded DNA generated. For each pair of primers, a standard curve was generated using 10-fold serially diluted gDNA extracted as in 2.5.1 against which the concentrations of unknown samples were determined. PCR and fluorescence detection were carried out in StepOne Real Time PCR system (Applied Biosystems). Analysis was performed in the accompanying software. The cycling conditions were as follows:

1. Activation of FastStart Taq DNA Polymerase. 10 min, 95°C
2. Denaturation. 15 sec, 95°C
3. Annealing and elongation. 1 min, 60°C

4. Steps 2-3 repeated 40 times

2.13.2 RNA stability assay

Overnight cultures were normalized to an OD_{600} of 0.003 in 25 ml of LB broth and grown for 3.5 h or 7 h to represent mid-exponential and a transition from exponential to stationary growth stages respectively. Degradation profiles of mRNA were studied by applying transcription inhibitor rifampicin at 500 µg/ml in 5 min intervals, extracting RNA, DNase treating and converting to cDNA as in 2.5.4, 2.12.1 and 2.12.3. To compare degradation rates of different mRNA species, cDNA amounts were quantified by RT-qPCR as in 2.13.1 using specific pairs of primers across untreated, 5-min, 10-min and 15-min post-rifampicin-treatment samples.

2.13.3 qPCR primer validation

The qPCR primers that were used for the *ihfA* and *ihfB* gene expression measurements, were validated to ensure that there was no cross-reactivity. That is the pair of primers designed to anneal to *ihfA* does not anneal to *ihfB* due to a degree of similarity in nucleotide composition between *ihfA* and *ihfB* and vice versa. To rule this out, qPCRs of the single IHF subunit knockout strains were carried out using the sets of primers for both *ihfA* and *ihfB.*as in 2.13.1.

2.14 Western immunoblotting

2.14.1 Protein lysate preparation for immunoblotting

An overnight culture was normalized to an OD_{600} of 0.003 in 25 ml of LB broth and grown to the required growth stage, typically early exponential, mid-exponential, transition from exponential to stationary and early stationary corresponding to 2 h, 3.5 h, 5 h and 7 h timepoints respectively. 2 OD_{600} units were collected by centrifugation at 3220 x g, 10 min, 4°C or 16000 x g, 1 min, 4°C depending on the volume, resuspended in 350 µl of PBS and transferred to

a sonication tube. The sample was sonicated on ice with 10 rounds of sonication pulses at an amplitude of 10 µm for 10 s with 30 s resting periods between. The resulting lysate was transferred into a clean Eppendorf tube and stored at -20°C.

2.14.2 Bradford assay

Protein concentration in whole cell lysates was determined by the Bradford assay to ensure equal loading on a gel. The assay is based on a red-to-blue colour change of Coomassie Brilliant Blue G-250 dye upon binding protein. The colour change is proportional to the amount of protein in a sample and can be measured spectrophotometrically at an absorbance of 595 nm. A 96-well plate was filled with 100 µl of 2x Bradford reagent and mixed with 100 µl of protein lysate samples diluted in PBS and the plate was read between 10 min and 60 min from mixing. A_{595} of the diluted samples were compared against a standard curve that was generated by serially diluting known amount of bovine serum albumin (BSA) in PBS.

2.14.3 SDS-PAGE gel assembly

Proteins were separated according to their mass by sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE). Briefly, denatured proteins bind to SDS in a gel that binds to them equally and equalizes their charge densities. This allows proteins to be separated according to their mass but not charge. First, proteins migrate through large pores of 5% acrylamide stacking gel at a low voltage causing them to concentrate in a straight narrow band. Then, they are separated by migrating through narrow pores of 12.5% acrylamide separating gel at a high voltage. pH gradient between the stacking and the separating gels leads to the stacking effect (Laemmli, 1970).

0.75 mm plates were cleaned and assembled for pouring the gel. Separating gel (12.5% acrylamide (National Diagnostics, Atlanta, US), 375 mM Tris-HCl pH 8.8, 0.001% SDS, 0.001% APS, 0.0001% TEMED) was poured first and overlaid with isopropanol to aid polymerisation. When polymerised, isopropanol

was poured off, rinsed with water and separating gel (5% acrylamide, 82.5 mM Tris-HCl pH 6.8, 0.001% SDS, 0.001% APS, 0.0001% TEMED) was poured and comb was added. When polymerised, the gel was assembled in the electrophoresis apparatus (Mini-PROTEAN system, BioRad, Hercules, US) and 1x running buffer (25 mM Tris base, 190 mM Glycine, 0.1% SDS) was added.

2.14.4 Loading and running SDS-PAGE gel

15 µl of 2x Laemmli buffer (4% SDS, 10% BME, 20% glycerol, 0.004% bromophenol blue, 0.125M Tris-HCl pH 6.8) was mixed with normalized amount of protein lysate according to results of the Bradford assay, as in 2.14.2, in a 30 µl total volume in 0.5 ml PCR tubes. The samples were denatured by boiling at 100°C for 5 min. 15 µl of the samples were loaded into each well with long tips, 3 µl of the broad-range prestained protein ladder (NEB) was used as a molecular marker, 5 µl of the Laemmli buffer was added into the empty wells. The samples were electrophoresed at 50 V until eluted from the wells and dye front linearized, and then at 150 V until the dye front was at the bottom of the gel.

2.14.5 Transfer of protein to PVDF membrane

To enable immunoblotting, proteins were transferred from the SDS-PAGE gel onto a polyvinylidene difluoride (PVDF) membrane with 0.2 µM pore size (Roche, Mannheim, Germany) by wet transfer. 1 litre of 1x transfer buffer (25 mM Tris base, 190 mM Glycine (Acros Organics, Geel, Belgium) with 20% methanol) was cooled down in a fridge to 4°C. PVDF membrane was cut to the gel size, activated in methanol for 1 min and soaked in transfer buffer for 10 min. 6 pieces of Whatman 5 filter paper were cut to the gel size and soaked in transfer buffer for 10 min. After electrophoresis was completed, the SDS-PAGE gel was briefly soaked in transfer buffer. The blotting sandwich was assembled in the following order from the cathode: 3 pieces of filter paper, gel, membrane, three pieces of filter paper. The blotting sandwich was secured in an electrophoresis tank, placed on ice and electrophoresed for 2 h at 300 mA.

2.14.6 Coomassie blue staining of SDS-PAGE gels

Staining of separated proteins on the SDS-PAGE gel is a way to check retrospectively if sample loading was equal and to compare protein profiles between samples. The gel was rinsed briefly in distilled water and stained in Coomassie blue (0.1% Coomassie Brilliant Blue R-250, 50% methanol, 10% glacial acetic acid) for 1 h, gently rocking. After the stain was removed, the gel was destained in 10% acetic acid for 10 min. The fresh 10% acetic acid was added, and the gel was destained overnight. When the staining intensity was deemed optimal, the gel was photographed.

2.14.7 Ponceau S staining of the membrane

To ensure that efficient transfer of proteins to membrane had occurred before proceeding with immunoblotting steps and to ensure equal loading, the PVDF membrane was stained with Ponceau S stain – negatively charged red stain that reversibly binds to positive and non-polar regions of proteins. This activity does not affect properties of membrane-bound proteins and thus further steps in Western immunoblotting. The membrane was flooded with Ponceau S solution, incubated for 5 min gently rocking, rinsed briefly with distilled water to wash off excess stain and viewed. The membrane was photographed. To destain, the membrane was flooded with 0.1 N NaOH and rinsed further with water.

2.14.8 Blocking and probing the membrane with protein-specific antibodies

To probe for a specific antigen among proteins present on a blotting membrane, the membrane was incubated with a primary antibody specific for that antigen. To visualise an antigen-primary antibody complex, the membrane was incubated with a secondary antibody that is specific for the primary antibody and in turn visualised using a chemiluminescent assay.

The membrane was blocked with blocking buffer (5% skimmed milk powder, PBST (PBS, 0.01% Tween-20 (Acros Organics, Geel, Belgium))) for 1 h, gently rocking. Experimental primary antibody and anti-DnaK antibody were diluted in the blocking buffer. Primary experimental antibodies – mouse anti-FLAG and mouse anti-Myc were used at 1:20000 dilution, mouse anti-H-NS was used at 1:1000, mouse anti-RpoD (Neoclone, Madison, US) was used at 1:20000 and were reused multiple times. Primary mouse anti-DnaK (Enzo Life Sciences, New York, US) that served as a loading standard was used at 1:100000 and was added freshly each time. The membrane was incubated with the primary antibody mix for 16 h.

After the primary antibody was decanted, the membrane was washed in PBST 3 times for 2 min each. Horseradish peroxidase (HRP)-conjugated goat anti-mouse secondary antibody (Bio-Rad, Hercules, US) was diluted 1:5000 in blocking buffer and incubated with the membrane for 90 min gently rocking. The membrane was washed in PBST 3 times for 2 min each, followed by one 2 min wash in PBS.

2.14.9 Visualisation of the chemiluminescent signal

Enhanced chemiluminescence (ECL) components luminol and peroxide formulations (Pierce^{TM}, Fisher, Loughborough, UK) were mixed 1:1. The ECL solution was applied to the membrane and it was sealed between plastic sheets. The reaction was allowed to proceed for 5 min in the dark, after which the chemiluminescent signal from the protein bands was visualised on the ImageQuant LAS 4000 imager. ImageJ software (available to download at imagej.nih.gov) was used to quantify protein bands by densitometry.

2.15 Mass spectrometry (MS) sample preparation

2.15.1 Protein lysate preparation for MS

An overnight culture was normalized to an OD_{600} of 0.003 in 25 ml of LB broth and grown to the required growth stage. 2 OD units were centrifuged to collect

cells at 3220 x g, 10 min, 4°C. The cell pellet was resuspended in 600 µl of 8 M urea and moved into a sonication tube. The sample was sonicated on ice with 10 rounds of sonication pulses at an amplitude of 10 µm for 20 s with 20 s resting periods between. Resulting lysate was transferred into a fresh Eppendorf tube and centrifuged at 1500 x g for 15 min at 4°C to remove cell debris. 400 µl of a supernatant was transferred into a fresh tube and centrifuged at 9300 x g for 10 min at 4°C. 200 µl of a supernatant was transferred into a fresh tube and kept at -20°C. The protein concentration was determined by Bradford assay as in 2.14.2 using 8 M urea (BDH, Poole, England) for dilutions of samples and BSA in a standard curve.

2.15.2 Protein tertiary structure removal and trypsin digestion

To reduce disulphide bridges, 100 µg of protein in 50 µl total volume was treated with 5 mM of dithiothreitol (DTT) shaking at 350 rpm for 10 min at 60°C. To prevent cysteine bridges re-forming, exposed cysteine residues were alkylated by adding 10 mM of iodoacetamide (IAA) for 30 min in the dark. To quench the reaction, dilute 8 M urea and neutralize pH in order to enable trypsinisation, 150 µl of 200 mM of ammonium bicarbonate was added. 20 µl of the sample preparation was transferred into a fresh tube, 2 µg of trypsin (Roche, Mannheim, Germany) was added and incubated at 350 rpm for 18 h at 37°C. Trypsin digestion was stopped by adding 0.1% trifluoroacetic acid (TFA).

2.15.3 Protein purification

Protein samples for mass spectrometry were purified and concentrated using ZipTip technology (Merck, Millipore, Darmstadt, Germany). $ZipTip_{C\text{-}18}$ is a 10 µL pipette tip with a 0.6 µL bed of chromatography media fixed at its end. Proteins eluted from a ZipTip are free of contaminants and are ready for MS analysis. C-18 resin has binding capacity of 5 µg when used with saturating amount of analyte and is ideal for purifying proteins with small molecular weight.

ZipTip manipulations involve very slowly aspirating and discharging to avoid introducing bubbles. The ZipTip was wetted using 10 µl of 100% acetonitrile

(ACN) (Fisher, Loughborough, UK) two times and equilibrated with 10 µl of 0.1% TFA two times. A protein sample was centrifuged at 9300 x g for 5 minutes at 4°C to pellet insoluble protein aggregates, and 15 µl of a supernatant was carefully transferred into a fresh tube without disturbing bottom of the tube. The ZipTip was charged with 10 µl of the sample ten times and washed with 10 µl of the wash buffer (0.1% TFA and 5% methanol) five times. 10 µl of the elution buffer (0.1% TFA and 50% ACN) was pipetted up, left in the tip for 1 min and eluted into a fresh tube three times with a fresh buffer. The sample was freeze-dried, resuspended in 25 µl of the resuspension buffer (0.5% acetic acid, 2.5% ACN) and centrifuged at 16000 x g for 5 min at 4°C. 20 µl of the sample was transferred into a new tube, A_{220} was measured on the Nanodrop 1000 to check peak presence between 220 and 230 nm and the sample was sent for the mass spectrometry.

2.16 Mass spectrometry and analysis

2.16.1 MS run

Complex trypsin-digested peptide mix was first separated by 1 h high performance liquid chromatography (HPLC) gradient using reverse phase C18 columns on an Dionex Ultimate 3000 UPLC (ThermoFisher Scientific) followed by a run on Q-Exactive (ThermoFisher Scientific) mass spectrometer located at the Mass Spectrometry Core, Conway Institute, UCD, Ireland. Q-Exactive work principle briefly. After separation by HPLC, the sample was injected into the mass spectrometer where it is ionized using atmospheric pressure chemical ionisation (APCI), ions were separated by their mass passing through quadrupole mass filter. The quadrupole was followed by a C-trap that injects ions into an HCD cell where selected ions are dissociated into smaller fragments by collision-induced dissociation. The fragment ions were sent back to the C-trap from where they are ejected into an Orbitrap detector to be analysed by tandem mass spectrometry (MS/MS) (Michalski *et al.*, 2011).

2.16.2 MaxQuant analysis

Mass spectrometry output files in the RAW format were loaded into the MaxQuant (available for download at coxdocs.org, version 1.6.10.43). The RAW files were searched against the *Salmonella enterica* serovar Typhimurium str. SL1344 complete proteome sequence obtained from UniProt (proteome ID UP000008962) assessed in February 2020. Carbamidomethylation of cysteine was chosen as a variable modification, while oxidation of methionine and N-terminus acetylation were chosen as fixed modification. Trypsin digestion was selected with a maximum of two missed cleavages. Label Free Quantification (Fast LFQ) mode was selected in group-specific parameters and iBAQ quantification was added in global parameters. Other settings were default. The search was performed using the Andromeda peptide search engine integrated into the MaxQuant. In the MaxQuant output, the summary of results was viewed in the ProteinGroups.txt file located in the combined folder. Peptides used for protein quantification were set to razor and unique peptides.

2.16.3 Perseus analysis

To perform statistics on protein quantification data, the Perseus software (version 1.6.10.50) that is a companion software for MaxQuant was used. Potential contaminants, proteins only identified by site and reverse hits were filtered out. Proteins were annotated using *Salmonella enterica* serovar Typhimurium str. SL1344 annotations downloaded from annotations.perseus-framework.org. Multiple annotation types (GO terms, Pfam name, gene name, full Uniprot name) were used to ensure that most proteins are annotated. The LFQ intensity values were $\log_2$-transformed, and the samples were grouped to account for the experimental design. A protein was considered to be present for the purpose of the analysis if it was detected in at least two samples in any of the sample groups. Two-sided T-tests were performed using FDR = 0.05 and S0 = 1 on the entire dataset to discover differentially expressed proteins. Hierarchical clustering was performed on these proteins (the $\log_2$-transformed LFQ values were normalized by Z-score) to distinguish between up and downregulated proteins. For all other proteins, that included proteins that were exclusive to either sample group and non-differentially expressed proteins,

imputation was performed that replaced missing values from normal distribution. Two-sided T-tests were performed using FDR = 0.05 and S0=1 on these proteins. Kegg annotation of the differentially expressed proteins was added manually using Kegg mapper available at www.genome.jp/kegg/tool/conv_id.html to simplify categorisation. Lists of differentially expressed and unique proteins were sorted in the order of increasing q-value and saved using Excel.

2.17 Microscopy

Bacterial cultures were grown to exponential or late stationary growth phase. 1 OD_{600} unit was centrifuged at 16000 x g for 1 min to harvest cells and the pellet was washed 3 times with 100 µl of PBS. 10 µl of the cell suspension was applied to a glass slide, dried near Bunsen burner and heat fixed. The slide was flooded with crystal violet for 1 min and excess stain was removed. The specimen was viewed under 1000x magnification with oil immersion on the Nikon Optiphot microscope. Images were captured on the Nikon DS-Fi1 camera attached to the microscope.

2.18 Competitive fitness assays

Bacterial strains were competed against each other by growth in the same flask starting from 1:1 ratio. Derivatives of each strain to be competed housing a chloramphenicol resistance cassette (*cat*) within a non-expressed pseudogene *SL1483* were created by P22 generalized transduction, as in 2.9, to enable selection. Strains to be competed were pre-conditioned separately as 25 ml cultures for 24 h without antibiotics, then 10^5 cells of each strain were mixed in 25 ml of fresh LB broth and grown for another 24 h. Plating on chloramphenicol-containing plates and plates with no antibiotics for enumeration was done at T=0 h and T=24 h. Taking WT SL1344 Vs OrfSwap$^{ihfA\text{-}ihfB}$ competition as an example, SL1344 was competed against OrfSwap$^{ihfA\text{-}ihfB}$*_SL1483::cat* and, as a control, SL1344*_SL1483::cat* was competed against OrfSwap$^{ihfA\text{-}ihfB}$. Competitive fitness was calculated according to the formula:

f.i. = (ln (Nc(24)/Nc(0)))/(ln ((Nwt(24)/Nwt(0))),

where Nc(0) and Nc(24) are initial and final counts of a competitor and Nwt(0) and Nwt(24) are initial and final counts of the WT. Competitor is a strain other than the WT. F.i.<1 means that the competitor is less fit than the WT, f.i.>1 is opposite.

2.19 Motility assays

Assays were carried out precisely as described to achieve agreement between biological replicates. 0.3% LB agar was made, agar was melted in a 100 ml bottle in a Tyndall steamer for 50 min, allowed to cool in a 55°C water bath for 20 min, six plates were poured and left to dry near a flame for 25 mins. 1 µl of bacterial overnight culture was pipetted under the agar surface with two inoculums per plate. Plates were placed into a 37°C incubator not stacking on each other to ensure equal oxygen access. After five hours the diameters of the resulting swarm zones were measured, and the ratio of the WT Vs mutant was calculated.

2.20 SPI-1 and SPI-2 reporter assays

Salmonella pathogenicity island (SPI) activity was accessed by measuring the expression of *gfp* reporter gene fusions to promoters of *prgH* and *ssaG* to look at SPI-1 and SPI-2 expression, respectively. The *gfp* reporter fusions were transduced into each strain by P22 generalized transduction as in 2.10 and selected with chloramphenicol. 100 µl of overnight culture of the *gfp* reporter-carrying strain was diluted 1:100 in LB broth. Black 96 plate with transparent flat bottom was filled with 100 µl of the diluted culture in six technical replicates, negative controls were included. The plate was sealed with parafilm and incubated at 300 rpm, 37°C for 24 h in the Synergy H1 microplate reader (Biotek, Vermont, USA). Bacterial growth was measured at 600 nm and GFP fluorescence was read using 485.5 nm excitation frequency at 528 nm emission frequency, measurements were taken every 20 min. For minimal medium SPI reporter assays, the pre-conditioned culture was normalised to an OD_{600} of 0.03

in the medium of the required $MgCl_2$ concentration and the measurements proceeded as above.

2.21 Minimum inhibitory concentration (MIC) of antibiotics determination

MIC_{90} of antibiotics (a minimal concentration at which 90% of bacterial growth is inhibited) was found by serially diluting antibiotics and spectrophotometrically testing the ability of different dilutions to inhibit bacterial growth. On a 96-well plate, all wells (excluding column 12) were filled with 60 µl of sterile LB broth. 1 ml solutions of antibiotics to be tested were prepared at the highest desired concentration in LB. 300 µl of the prepared antibiotics were added to the wells of column 12 and homogenised by pipetting up and down 5 times with a multichannel pipette. 240 µl was transferred to the next wells in column 11, homogenisation was repeated and serial 1:1.25 dilutions were sequentially continued until column 3. The final 240 µl from column 3 were discarded. All the wells were inoculated with bacterial cultures normalized to an OD_{600} of 0.003 except column 1. In this way column 1 contained negative controls (no bacteria and no antibiotic), column 2 contained positive controls (no antibiotics) and columns 3-12 contained serially diluted antibiotics inoculated with the identical number of bacteria. The plate was covered, sealed between plastic sheets and incubated for 18 h at the standard growth conditions. The plate was read by measuring OD_{600} values on a plate reader (Multiscan EX, Thermo Electronics).

2.22 Cell culture

2.22.1 Macrophage entry and viability assays in SPI-2 inducing conditions

To assess entry and intracellular survival of *Salmonella* into RAW264.7 macrophages by phagocytosis, SPI-2 induced bacteria were used. To ensure SPI-2 induction and, therefore, uptake of *Salmonella* by phagocytosis, bacteria were grown to late stationary growth phase and complement-opsonised (Lee & Falkow, 1990). 10 ml of an overnight culture in 125 ml conical flask in LB broth

was set up and grown at the standard conditions for 18 h. 500 µl of the overnight culture was centrifuged at 16000 x g for 1 min and resuspended in 500 µl Hank's buffered saline solution ($HBSS^{-/-}$) (Corning, Manassas, US). Bacterial cells were opsonised by diluting 1:10 in DMEM and 15% normal mouse serum (NMS) and incubated for 30 min at 37°C. After opsonisation, bacteria were pelleted and resuspended in a corresponding amount of $HBSS^{-/-}$.

Gentamycin protection assay was used to determine bacterial counts inside macrophages. Monolayers were washed twice with 500 µl of $HBSS^{+/+}$. It was important to ensure that during all washes a monolayer stayed dry only briefly, that is as medium was aspirated, HBSS wash was added straight away well by well. Infection medium was made by mixing opsonised bacteria with DMEM+FBS supplemented with 500 µg/ml of L-Histidine (always supplemented into DMEM during infection for optimal growth of SL1344). Monolayers were infected with bacteria using MOI 20 in three technical replicates for each timepoint and strain. The plate was centrifuged at 200 x g for 10 min to synchronize infections and incubated for 30 min at 37°C, 5% CO_2. In the meantime, the infection medium was plated for enumeration on LB agar plates – T=0 h. To kill all extracellular bacteria, the monolayers were washed once with $HBSS^{+/+}$ and high gentamycin (100 µg/ml) treatment diluted in DMEM+FBS was added to the wells. The plate was incubated at 37°C, 5% CO_2 for 1 h. At 1 h post infection the monolayers were washed three times with $HBSS^{+/+}$, macrophages were lysed by adding 1 ml of ice-cold water, pipetting up and down ten times with scraping and intracellular bacteria were plated for enumeration. The monolayers which were intended for other timepoints, were washed once with $HBSS^{+/+}$, low gentamycin (10 µg/ml) treatment in DMEM+FBS was added and the plate was incubated at 37°C, 5% CO_2. The low gentamycin concentration is to ensure that any extracellular bacteria are killed, but at the same time to avoid gentamycin permeabilizing plasma membrane of a macrophage (Kaneko *et al.*, 2016). At later timepoints monolayers were washed three times with $HBSS^{+/+}$, macrophages were lysed by adding 1 ml ice-cold water, pipetting up and down ten times with scraping and intracellular bacteria were plated for enumeration.

2.22.2 Macrophage entry and viability assays in SPI-1 inducing conditions

To assess active entry of *Salmonella* into RAW264.7 macrophages and intracellular survival, SPI-1 induced bacteria was used. Bacteria were grown to mid-exponential growth phase in LB broth to maximize SPI-1 expression (Steele-Mortimer *et al.*, 1999). Overnight cultures were subcultured 1:33 in 10 ml of fresh LB broth in 125 ml conical flask and grown for 3.5 h. 500 µl of the culture was centrifuged at 16000 x g for 1 min and resuspended in 500 µl of $HBSS^{-/-}$. Monolayers were washed twice with 500 µl of $HBSS^{+/+}$ and infected with bacteria at MOI of 5. All the further manipulations were as in 2.22.1.

2.23 Global supercoiling determination

2.23.1 Reporter plasmid assays for bacteria grown in rich media

The global DNA supercoiling is defined by the linking number (number of times one DNA strand crosses trajectory of the other DNA strand in DNA molecule, *Lk*) that is composed of the twist and the writhe components. Changes to the *Lk* can be made one step at a time by breaking a single stand of the double helix (type I topoisomerases) or in two steps by introducing a dsDNA break, passing the other strand through the break and resealing it (type II topoisomerases) (Higgins & Vologodskii, 2015). The linking number set in the cytoplasm prior to extraction is preserved in a plasmid such as pUC18 and can be determined by analysing its electrophoretically separated topoisomers. In this analysis the extent of DNA supercoiling was judged by identifying the most prominent band in each sample and comparing its migration distance from the loading wells with that of the most prominent band in the neighbouring lanes. When the most prominent band was shifted towards the region of higher supercoiling density, negative Δ*Lk* value was assigned, when it was shifted towards the relaxed region, positive Δ*Lk* value was assigned.

Global DNA supercoiling was assayed in bacterial strains transformed with a reporter plasmid pUC18 as in 2.8. An overnight culture of pUC18-containg strain was normalized to an OD_{600} of 0.003 and grown to the mid-exponential

(3.5 h) or late stationary growth stage (24 h) in 25 ml LB broth. For the late stationary phase plasmid extractions, 14 OD_{600} units were harvested and pUC18 was isolated as in 2.4.2. For the mid-exponential plasmid extractions, the entire culture volume was split in three and three pUC18 isolations were performed in parallel to increase the plasmid yield. At the plasmid elution stage, the three fractions were pooled together.

To observe the range of DNA supercoiling states characteristic of a strain at a given growth stage, extracted pUC18 samples were resolved on 0.8% agarose gel supplemented with the DNA intercalating agent chloroquine. 2 L of 1x TBE buffer (89 mM Tris base, 89 mM boric acid, 2 mM EDTA pH 8.0) and 1 ml of 25 mg/ml chloroquine were made. 0.8% agarose solution was made from 300 µl TBE and melted in a Tyndall steamer. When the gel cooled down, it was supplemented with 2.5 µg/ml chloroquine. Chloroquine intercalation converts negatively supercoiled plasmids to open, relaxed ones. The 27 cm long gel was poured, left to polymerise for 2 h and covered with 1.7 litres of the running buffer containing 1x TBE and chloroquine at 2.5 µg/ml. 1 µg or 500 ng of the plasmid samples in 15 µl volumes were mixed with 5x loading dye (80% glycerol, 0.01% bromophenol blue) and loaded on a gel. The gel was electrophoresed for 16 h at 100 V. The gel was washed in distilled water for 24 h changing water a few times, stained in 1 µg/ml ethidium bromide for 1 h rocking in the dark. The stain was poured off and the gel was washed in distilled water for further 1 h. The plasmid topoisomers were visualised under UV on the ImageQuant LAS 4000 imager. ImageJ software was used to outline plasmid topoisomer distribution profiles.

2.23.2 Reporter plasmid assays for bacteria grown in minimal medium

To determine global DNA supercoiling of bacteria grown in N minimal medium with different Mg^{2+} concentrations, the pUC18-containg strains were cultured as in 2.2.2. 6 OD_{600} units were collected by centrifugation for plasmid extraction as in 2.5.2. The samples were electrophoresed on a chloroquine gel as in 2.23.1.

Chapter 3:

Repositioning of *ihfA* and *ihfB* in *Salmonella enterica* serovar Typhimurium and its phenotypic significance

3.1 Introduction

Since the advent of whole genome sequencing and as the technology has been getting cheaper and more high-throughput, researchers have been sequencing their favourite organisms so that thousands of bacterial genomes became available in genome repositories (Heather & Chain, 2016; Sayers *et al.*, 2019). This made possible comparative analysis between bacteria for multiple purposes such as improving microbial taxonomy (Glöckner *et al.*, 2017). Another application of vast genomic data is comprehension of chromosomal positions of genes in microbial species. This is of interest as it will help to determine why certain genes are positioned in certain loci on a chromosome and what advantage for bacteria does it confer. Deep understanding of gene arrangement principles will aid a field of synthetic biology to construct artificial chromosomes or to modify existing ones with the aim to yield an organism with enhanced useful properties.

Observing genomes revealed that the positions of genes encoding elements of the major regulatory pathways in the chromosomes of related species of bacteria are highly ordered and that this order is greatly conserved (Sobetzko *et al.*, 2012; Slager & Veening, 2016; Yubero & Poyatos, 2020). *Salmonella's* genetic arrangement is highly similar to that of *Escherichia coli* – a species that belongs to the same family (the *Enterobacteriaceae*) as *Salmonella*. For example, positions of the genes that encode master transcriptional regulators in *E. coli* and *Salmonella* are almost identical. These include genes that encode nucleoid associated proteins (NAPs), DNA topoisomerases, and RNA polymerase modulators. The conservation can be traced in terms of gene order, location within replicores, macrodomain positioning and relative distance from the origin of chromosomal replication (Sobetzko *et al.*, 2012). While distance from the origin seems to be a highly conserved property, relative inversions over the terminus region are frequent and well-tolerated (Alokam *et al.*, 2002).

One way to test if the position of a certain gene is important is to alter its location on a chromosome and to test if properties of the resulting mutant differ from those of the wild type. This approach was used in a number of studies. To assess the dependence of promoter activity of location, NAP H-NS promoter-

reporter fusion P*hns-yfp* was placed into various locations on the chromosome and expression changes were monitored, showing strong dependence of the promoter activity on location (Brambilla & Sclavi, 2015). The whole *dusB-fis* operon encoding another NAP – FIS was transferred into a new location in the other study. This affected not only *fis* expression, but also impaired competitive fitness of the cell and various stress responses (Gerganova *et al.*, 2015). The open reading frames of *hns* and *stpA*, that encode paralogous NAPs, were reciprocally exchanged in *Salmonella* while keeping the original 5' and 3' regulatory regions in place (Fitzgerald *et al.*, 2015). The resulting gene expression patterns of the relocated genes and production patterns of the corresponding proteins did not simply acquire characteristics of each other, highlighting the complexity within the bacterial regulatory networks. An unexpected finding was that the resulting strain had competitive fitness that was superior to that of the wild type at 25°C (Fitzgerald *et al.*, 2015). Whole gene exchange and ORF exchanges of *fis* and *dps* was performed in another recent study. These gene encode NAPs that are expressed during different stages of the growth cycle and possess many opposing properties. It was found that the gene exchange had much less consequences than the ORF exchange in terms of global gene expression and global Fis binding throughout the genome (Bogue *et al.*, 2020). The present study aimed to reposition ORFs of the genes that encode subunits of another NAP – IHF.

IHF (integration host factor) is in the centre of this study. It was discovered as having an essential role in site-specific integration and excision of bacteriophage λ (Miller & Friedman, 1980). Since then, the roles of IHF in site-specific recombination, replication and transcription, transposition and genomic architecture were described (Dorman, 2014). All these functions are mediated via IHF's unique way of interacting with DNA. The structure of IHF was solved when it was co-crystallised with DNA to determine that it binds DNA as a single heterodimer and introduces a sharp 180° bend that explains its architectural importance (Rice *et al.*, 1996). This bending has a potential to bring together distant DNA elements, thus justifying the gene regulatory importance of IHF. IHF structure is highly similar to its paralogous NAP HU, but fundamentally

differs from it in that IHF binds DNA in a sequence-specific manner (Swinger & Rice, 2004).

In *Salmonella* IHF is encoded by the genes *ihfA* and *ihfB* that are dispersed on the chromosome (Fig. 1.1). These positions are conserved in *E. coli*, except that a Ter-centred inversion has moved *ihfA* to the opposite replicore while maintaining a similar distance from the origin. Since different members of gamma-proteobacteria evolved to retain many of their regulatory genes at similar genomic locations, it seems plausible that this must confer some selective advantage. The advantage of having genes that encode subunits of the same protein located at a distance from one another is not obvious.

The *ihfA* and *ihfB* transcription units are aligned with the movement of the replication fork (Fig. 1.1). This is in accordance with the bias to co-orientation of transcription and replication to avoid steric collision between RNA polymerase and DNA polymerase that runs chromosome replication bidirectionally from the origin (Srivatsan *et al.*, 2010). However, the rationale for other aspects of positioning of these genes is less evident. Not only is IHF encoded by two genes in two different loci, their regulation is independent (Aviv *et al.*, 1994). IHF can bind DNA as a homodimer *in vitro* (Zulianello *et al.*, 1994), but *in vivo* homodimers have not been found. However, their existence cannot be ruled out because *ihfA* and *ihfB* individual deletion mutants are viable and have transcriptomes that differ from those of a mutant with both IHF subunit genes deleted (Mangan *et al.*, 2006). Thus, it might be an advantage for these genes to be co-regulated and expressed from the same operon.

The Lambda Red-mediated recombination system was extensively used for strain construction (Datsenko & Wanner, 2000). Normally, most bacteria cannot be transformed with linear exogenous DNA as it gets efficiently eliminated by bacterial restriction-modification (R-M) systems or intracellular exonucleases (Bickle & Krüger, 1993; Lovett, 2011). One way to aid such transformation is to use strains deficient in some of the R-M components or exonucleases (Bullas & Ryu, 1983), however, this is not a universal method. Bacteriophages use their own recombination systems to ensure incorporation of their genetic material into bacterial DNA. For example, lambda phage carries the Red system that is

composed of three genes: *γ*, *β* and *exo*, that code for Gam, Bet and Exo protein products respectively. Bet and Exo promote homologous recombination, while Gam inhibits the host RecBCD exonuclease V (Murphy, 1998). Plasmid pKD46 carries the Lambda Red system under the control of an arabinose-inducible promoter. The replication of plasmid pKD46 is temperature-sensitive, allowing to be eliminated by growth at the non-permissive temperature once the recombination process is completed (Murphy, 1998).

To test the importance of the evolutionary conserved positions of IHF-encoding genes, their open reading frames were exchanged. Following the relocation, *ihfA* ORF became flanked at its 5' end by the P_{ihfB} promoter and associated regulatory 5' sites and at its 3' end, by the transcription terminator that originally flanked *ihfB*. The reciprocal situation applied in the case of *ihfB* ORF. Combination of the two relocations resulted in the OrfSwap$^{ihfA\text{-}ihfB}$ strain. A plan to exchange entire *ihfA* and *ihfB* genes including their respective regulatory regions was abandoned because each gene is a part of a larger, complex operon (Weglenska *et al.*, 1996; Wertheimer *et al.*, 1988). This complexity made it too difficult to define the true extent of each regulatory region with sufficient precision. Moreover, the genetic rearrangements would have affected the expression of the other genes in these operons. Instead, the effects of IHF subunit co-expression from the same genomic location was tested by engineering bicistronic operons. The ORF of *ihfA* was positioned directly downstream of *ihfB* under the control of the P_{ihfA} and *vice versa* for *ihfB*. These manipulations created artificial *ihfBA* and *ihfAB* operon (Table 2.1). The construction of these strains as well as phenotypic and functional outcomes of the repositions are described in detail below.

3.2 Results

3.2.1 Construction of the Δ*ihfA*, the Δ*ihfB*, the OrfSwap$^{ihfA\text{-}ihfB}$, the *ihfAB* operon and the *ihfBA* operon strains

To create OrfSwap$^{ihfA\text{-}ihfB}$, intermediate 2x*ihfA* and 2x*ihfB* strains were constructed first that contained, respectively, two copies of the *ihfA* gene and two copies of the *ihfB* gene (Table 2.1). Positioning *ihfA* in place of *ihfB* to construct 2x*ihfA* is schematically shown in Fig. 3.1 to represent the general strategy employed throughout this project. Briefly, a gene to be moved was tagged with an antibiotic resistance gene using plasmid pKD3 (carrying a chloramphenicol resistance cassette) or plasmid pKD4 (carrying a kanamycin resistance cassette) as a template. For 2x*ihfA* construction, a pair of primers (infA.cmR.Pfwd and infA.cmR.Prev; Table 2.2) was used to amplify a chloramphenicol resistance cassette, producing an amplimer that had overhangs homologous to a region immediately downstream of the gene of interest (*ihfA*). All insertions of the antibiotic resistance cassettes and amplicons in the subsequent steps were typically made three nucleotides downstream of the stop codon to avoid interrupting a terminator and excluding other regulatory regions. The second overhang was designed so as to delete 2-4 nucleotides from the target region to make insertion more efficient. PCR was carried out using high-fidelity Phusion DNA Polymerase. Resulting amplicon was purified with the Qiagen PCR purification kit and transformed by electroporation into a strain harbouring pKD46. This strain was grown with arabinose at 30°C to drive the expression of the Lambda Red system encoded on pKD46 (Datsenko & Wanner, 2000). The integrity of successful transformant colonies was confirmed by PCR with “confirmation” primers and sequenced by Sanger sequencing. Taking 2x*ihfA* as an example, the *ihfA*::Cm^R construct was similarly amplified with primers that had overhangs homologous to regions upstream and downstream of the *ihfB* ORF (infB.int.Pfwd and infB.int.Prev; Table 2.2). The amplicon was purified and transformed into the WT strain harbouring pKD46, successful transformations were confirmed as described above. Alongside the 2x*ihfA*, a 2x*ihfB* strain was constructed using a similar approach but using a kanamycin resistance cassette for selection. A P22 lysate of the 2x*ihfB* was used to transduce the *ihfB*::kan^R into the 2x*ihfA* background to yield the

OrfSwap$^{ihfA\text{-}ihfB}$. Antibiotic resistance cassettes were removed by the FLP-mediated site-specific recombination (Cherepanov & Wackernagel, 1995). The resulting strain had its *ihfA* ORF under the control of the regulatory region of *ihfB* and the *ihfB* ORF under the control of the regulatory region of *ihfA*. (Table 2.1)

To create *ihfAB* and *ihfBA* strains, *ihfA* and *ihfB* ORFs tagged with the Cm^R and Kan^R cassettes created during the OrfSwap$^{ihfA\text{-}ihfB}$ construction were used. The *ihfA*::Cm^R construct including 25 nucleotides upstream of the start codon was amplified with primers that had both overhangs homologous to sequence immediately downstream of *ihfB* (ihfB.int.ihfA::Cm_Pf and ihfB.int.ihfA::Cm_Prev; Table 2.2). This was done to ensure that the ribosome binding site of *ihfA* was brought with the ORF to enable translation of IhfA protein from *ihfBA* bicistronic mRNA. The amplicon was purified and transformed into the WT strain harbouring pKD46. The transformants were screened for the correct insertion, the Cm^R cassette was removed and the resulting *ihfBA* operon was sequenced as described above. The *ihfAB* operon strain was made using a protocol similar to that used for the *ihfBA* strain. The resulting *ihfAB* strain had two copies of *ihfB*, while *ihfBA* strain had two copies of *ihfA.* The second, stand-alone, copies of these genes were removed from their native loci by a method based on the λ-Red-mediated homologous recombination as described below. The final strains had *ihfAB* and *ihfBA* bicistronic operons, where transcription was driven by the regulatory input of *ihfA* and *ihfB*, respectively (Table 2.1).

To create the deletion strains with Δ*ihfA* and Δ*ihfB* mutations*,* the λ-Red recombination system was used to insert a kanamycin resistance cassette in-frame between the start codon and last 21 bp of the target *ihf* gene (Baba *et al.*, 2006). The insertion was made in-frame to avoid polar effects on downstream genes when the resistance cassette was removed. The same method was used to delete a second copy of a gene from its original locus in the *ihfAB* and *ihfBA* operon strains.

IHF plays a role in tyrosine-integrase-mediated site-specific recombination (Miller & Friedman, 1980; Dorman & Bogue, 2016). Therefore, attempts to

remove the antibiotic resistance cassettes flanked by FRT sites by FLP-integrase-mediated site-specific recombination, in strains where IHF heterodimer is not present, such as 2x*ihfA*, 2x*ihfB*, Δ*ihfA* and Δ*ihfB*, failed. Hence, it should be noted that the final versions of these strains contain antibiotic resistance cassettes, while the cassettes were removed from all other strains to avoid potential fitness costs associated with carrying those cassettes.

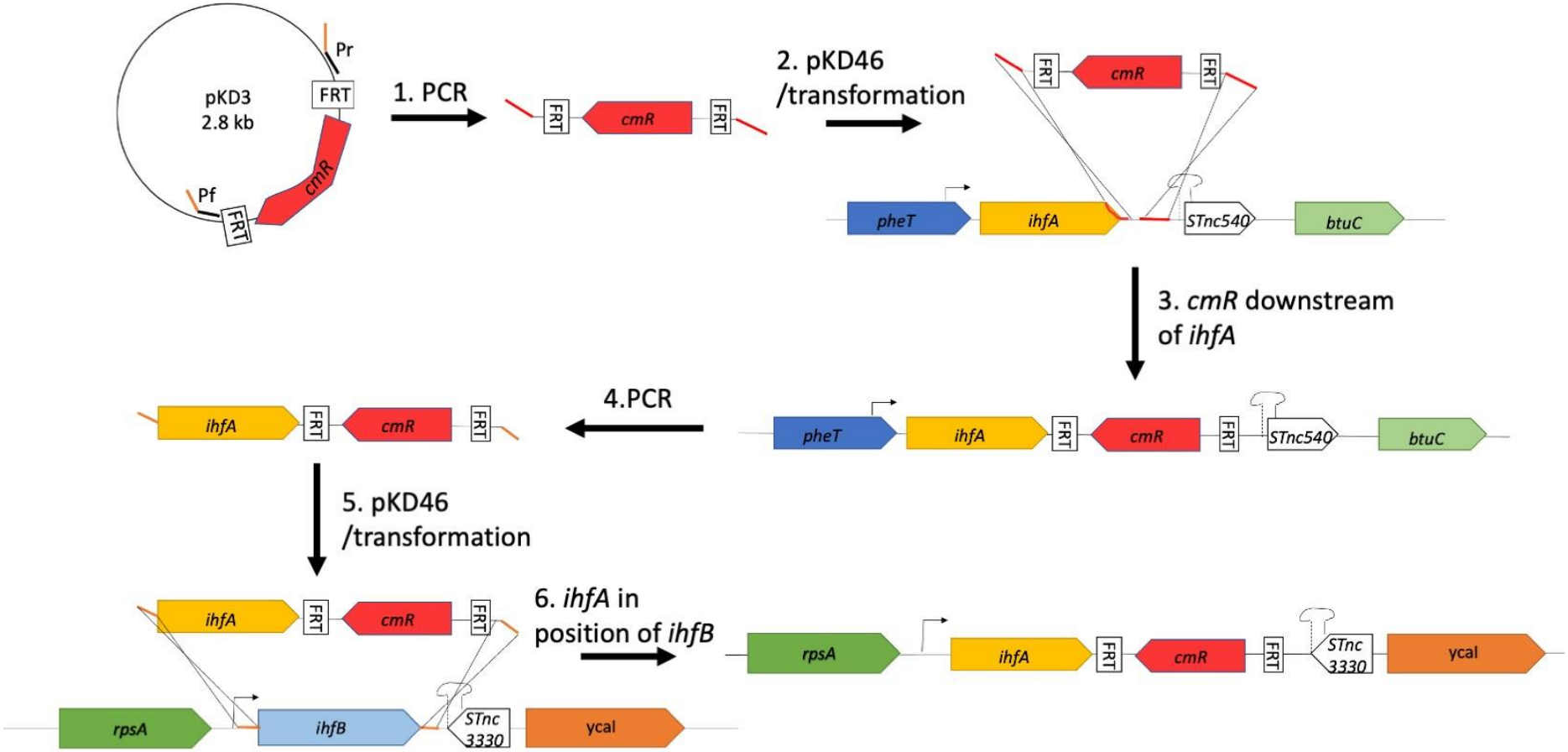

pKD3
2.8 kb
Pr
Pf
FRT
cmR
1. PCR
2. pKD46
/transformation
pheT
ihfA
STnc540
btuC
3. cmR downstream
of ihfA
4.PCR
5. pKD46
/transformation
6. ihfA in
position of ihfB
rpsA
ihfB
STnc
3330
ycal

Figure 3.1 Strain construction strategy. An example of placing *ihfA* in position of *ihfB* (2x*ihfA* construction). 1. Chloramphenicol resistance cassette was amplified by PCR from the carrier pKD3 plasmid with a pair of primers that had overhangs (depicted in red) homologous to the region downstream of *ihfA*. 2. Linear PCR product was purified and transformed into the WT SL1344 harbouring pKD46. 3. In number of cases the linear product was inserted into the target region by the Lambda Red-mediated recombination yielding intermediate strain with Cm^R-tagged *ihfA*. 4. In another round of PCR, the *ihfA*-Cm^R construct was amplified with a pair of primers that had overhangs homologous to the regions just upstream and downstream of *ihfB*. 5. Linear PCR product was purified and transformed into the WT SL1344 harbouring pKD46. 6. In number of cases the linear product was inserted into the target region by the Lambda Red-mediated recombination yielding a strain with *ihfA*-Cm^R in place of *ihfB*. Resistance cassette was then removed by pCP20-mediated FLP site-specific recombination where needed.

3.2.2 Strain validation by whole genome sequencing

Manipulating with genes that encode pleiotropic global regulators such as IHF, may lead to compensatory mutations arising elsewhere in the genome. In addition, the strain construction strategy contains multiple steps and it would not be surprising if sequence errors, that arose at one of the steps, were carried down the construction pipeline. To ensure that any phenotypes observed in the strains OrfSwap$^{ihfA\text{-}ihfB}$, *ihfAB* and *ihfBA* are solely due to the intended chromosomal repositionings, entire genomes of these strains were sequenced at the Sanger institute (Hinxton, Cambridgeshire, UK). Sequence contigs were assembled guided by the reference SL1344 sequence NC_016810.1 using *de novo* assembly with Velvet (Zerbino, 2010) and any SNPs were found by alignment of the assembled genome with the same reference using Breseq (Deatherage & Barrick, 2014). New junctions were detected at *ihfA* and *ihfB* loci in all sequenced strains as expected, confirming successful gene rearrangements. The bioinformatic analysis was kindly performed by Dr. Aalap Mogre.

Two SNPs were detected in all of the sequenced strains. These were not a cause of concern, because the same mutations were found in the parental SL1344 WT strain and are present in all of the SL1344-derived strains in our lab (Fitzgerald *et al.*, 2015). These SNPs are: TTA→TTG resulting in L148L silent change in *menC* – encoding O-succinylbenzoate synthase and GAG→GTG resulting in a E95V amino acid substitution in the *manX* open reading frame – encoding a putative mannose specific permease. Apart from these previously described changes, no other SNPs were detected in the OrfSwap$^{ihfA\text{-}ihfB}$, *ihfAB* or *ihfBA*. Therefore, it was safe to conclude that any results observed when working with these strains were due to the actual chromosomal repositionings of the *ihfA* and *ihfB* ORFs.

3.2.3 Growth characteristics of strains with repositioned *ihf* genes

To assess the impact of the repositionings of the IHF-encoding genes on the growth characteristics of *Salmonella*, the growth profiles of OrfSwap$^{ihfA\text{-}ihfB}$, *ihfAB* and *ihfBA* were monitored in liquid medium by optical density at 600 nm (OD_{600}) and compared to the WT. To encompass all growth phases, OD_{600} measurements were taken every hour until 3 h, then every 30 min until 8 h and lastly at 24 h. None of the strains with repositioned *ihf* genes differed in their growth pattern from the WT (Fig. 3.2a and Fig. 3.2c). Strains 2x*ihfA* and 2x*ihfB*, the intermediate strains for the OrfSwap$^{ihfA\text{-}ihfB}$ construction, grew slower than the WT at late stationary growth phase (Fig. 3.2b). Growth of the Δ*ihfA* and Δ*ihfB* strains was shown previously to be indistinguishable from the WT (Mangan *et al.*, 2006), and this was confirmed here (Fig. 3.2d).

The growth curves obtained were used to estimate the timing of various growth stages and growth stage transitions. Thus, the 2 h, 3.5 h, 5 h, 7 h and 24 h timepoints were chosen to represent lag, mid-exponential, exponential-stationary transition, early stationary and late stationary phases of bacterial growth. Therefore, to harvest bacteria at the desired growth phase in further experiments, cultures were grown in the standard conditions to the desired reference timepoint before harvesting a certain amount of OD_{600} units.

OD_{600} measurements give a good indication of bacterial growth patterns, however, differences in cell morphology or size can affect light scattering during optical density measurements and render comparisons between strains invalid. In addition, this method does not distinguish live from dead cells. Therefore, growth was also measured by calculating the number of colony-forming units in a liquid culture by spreading aliquots onto agar plates every 2 h until the 8-h and at the 24-h. There was no viable count difference between the WT and the strains with repositioned *ihf* genes (Fig. 3.3a, Fig. 3.3c). Viable counts of strains 2x*ihfA* and 2x*ihfB*, the intermediate strains for the OrfSwap$^{ihfA\text{-}ihfB}$ construction, were lower than those of the WT from 4 h onwards (Fig. 3.3b). The Δ*ihfA* and Δ*ihfB* strains' viable counts were lower than that of the WT only at 6 h (Fig. 3.3d). Together, these results suggest that in the strains where *ihfA* and *ihfB* are repositioned but both ORFs are present, the growth cycle is not affected.

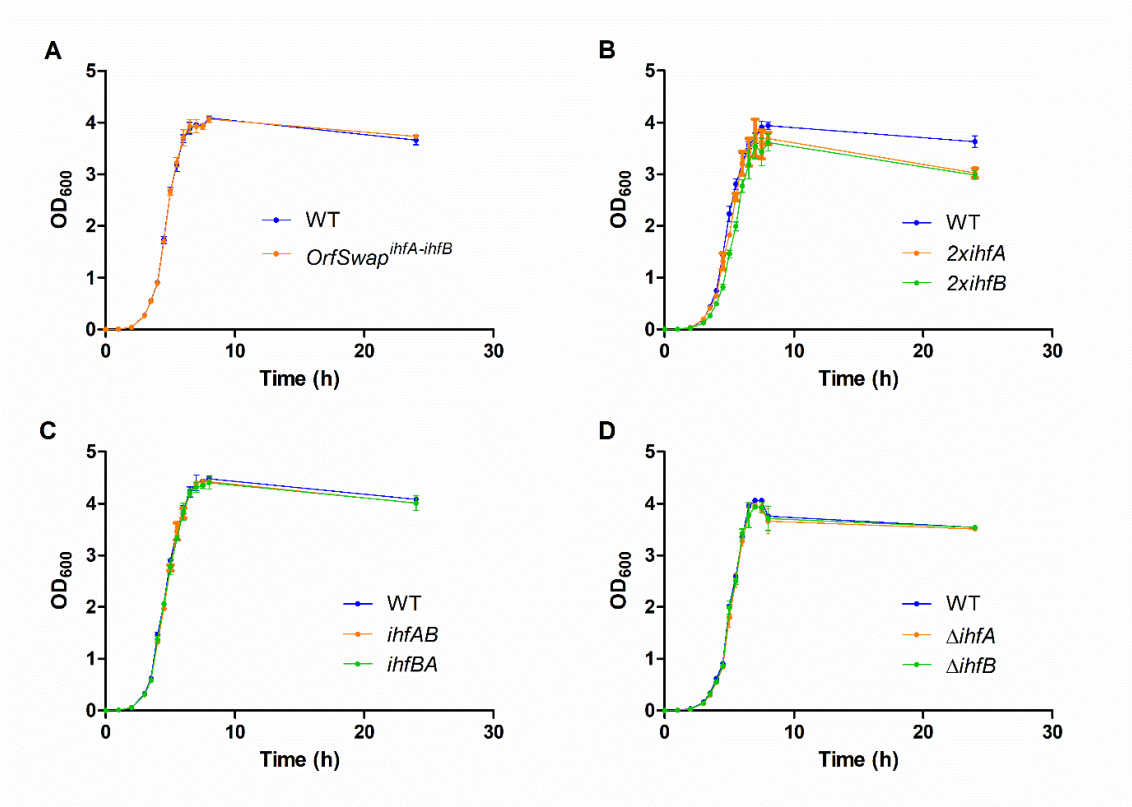


Figure 3.2 Growth curves of the strains used in the IHF work as measured by absorbance at 600 nm. OD_{600} values were taken every hour until 3 h, then every 30 min until 8 h and lastly at 24 h. a) Growth patterns of the OrfSwap$^{ihfA\text{-}ihfB}$, as compared against the WT. b) Growth patterns of the 2x*ihfA* and the 2x*ihfB* intermediate strains, as compared against the WT. c) Growth patterns of the IHF operon strains *ihfAB* and *ihfBA*, as compared against the WT. d) Growth patterns of the IHF individual subunit deletion strains Δ*ihfA* and Δ*ihfB*, as compared against the WT. All plots are results of at least three biological replicates, error bars represent standard deviation.

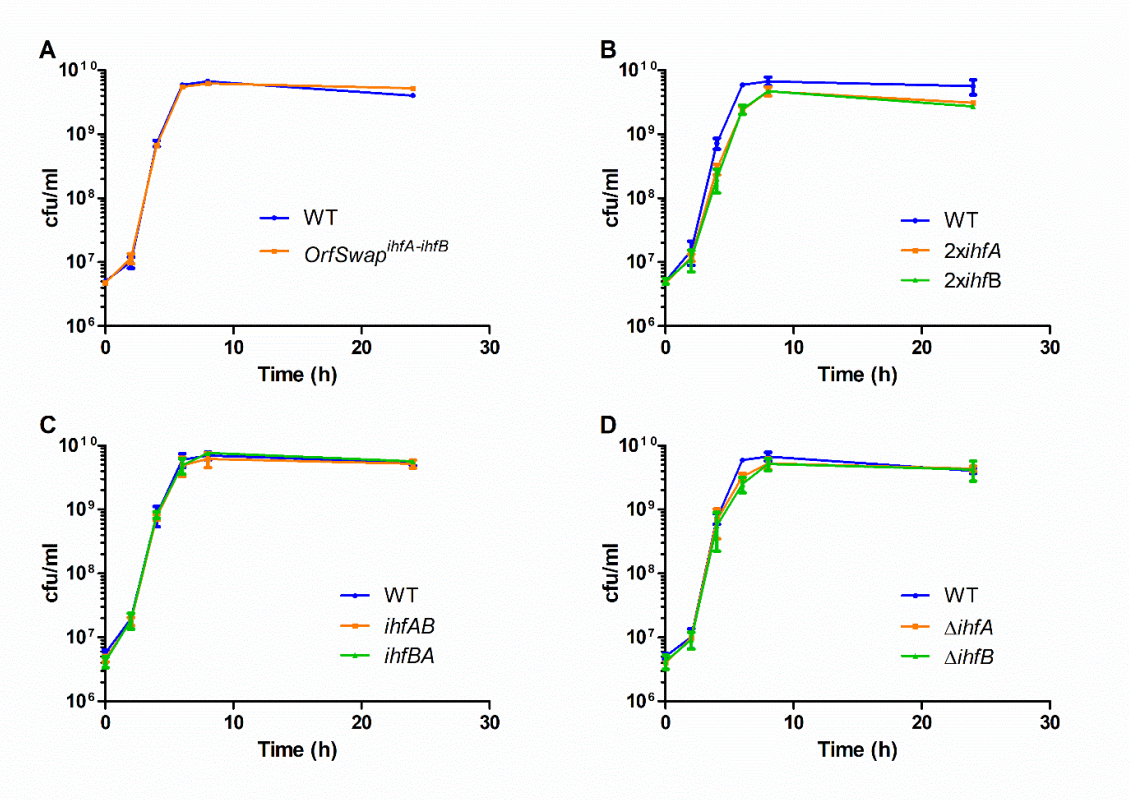


Figure 3.3 Growth curves of the strains used in the IHF work as measured by viability counts. Viability counts were measured by spreading dilutions of bacterial cultures onto agar plates, incubating at 37°C and counting colonies. a) Growth patterns of the OrfSwap$^{ihfA\text{-}ihfB}$, as compared against the WT. b) Growth patterns of the 2x*ihfA* and 2x*ihfB* intermediate strains. Both strains grew significantly slower than the WT at 4 h, 6 h, 8 h and 24 h. c) Growth patterns of the IHF operon strains *ihfAB* and *ihfBA*, as compared against the WT. d) Growth patterns of the IHF individual subunit deletion strains Δ*ihfA* and Δ*ihfB*, as compared against the WT. Both Δ*ihfA* and Δ*ihfB* grew significantly slower than the WT at 6 h. All plots are results of at least three biological replicates, error bars represent standard deviation. One-way ANOVA with Tukey's Multiple Comparison Test was used to determine significance, where $p<0.05$.

3.2.4 Cell morphologies of strains with repositioned *ihf* genes

Each of the strains with the repositioned IHF-encoding genes, the *ihf* deletion strains and the WT were visualised by phase-contrast microscopy at the mid-exponential and the late stationary phases of growth. *Salmonella* is a short, rod-shaped cell and any changes to this morphology may be indicative of problems with the cell cycle. The cell cycle consists of a series of complex events whose main stages are chromosome replication, segregation and cell division (Reyes-Lamothe *et al.*, 2012). IHF is known to have a role in the initiation of chromosomal replication through the binding at *oriC*, the origin of replication (Ryan *et al.*, 2004). Therefore, strains where IHF is not fully functional may show morphological aberrations, such as cell filamentation.

No changes in cellular morphology were observed in the OrfSwap$^{ihfA\text{-}ihfB}$, *ihfAB* or *ihfBA* strains in either mid-exponential or late stationery growth phases when compared to the WT (Fig. 3.4 and Fig. 3.5). This suggested that despite repositioning of the *ihfA* and *ihfB* ORFs under the control of the regulatory regions of each other's gene, IHF retained its functionality in the cell cycle and that the assumptions about the location of ribosome binding sites for the downstream genes made when constructing the bicistronic *ihf* operons were justified. Strains where only one of the IHF subunits was present – 2x*ihfA*, 2x*ihfB*, Δ*ihfA* and Δ*ihfB,* displayed filamentous morphologies during both growth phases (Fig. 3.4 and Fig. 3.5), confirming the importance of heterodimeric IHF during the cell cycle. Note that the cell filaments visible in figures 3.4 and 3.5 do not allow to conclude that one strain forms more filaments than the other as no comparative filament analysis across different fields of view was performed.

The filamentous phenotypes of the Δ*ihfA* and Δ*ihfB* strains provide explanation for the deficiency in growth seen at 6 h when measured by viable counts but not by absorbance at 600_{nm} (Fig. 3.2d and Fig. 3.3d). This is because cells with a morphology different from that of the WT scatter light in a different way. Thus, OD_{600} measurements misleadingly show that amount of bacterial growth is the same as in the WT, although there are more viable cells in the WT.

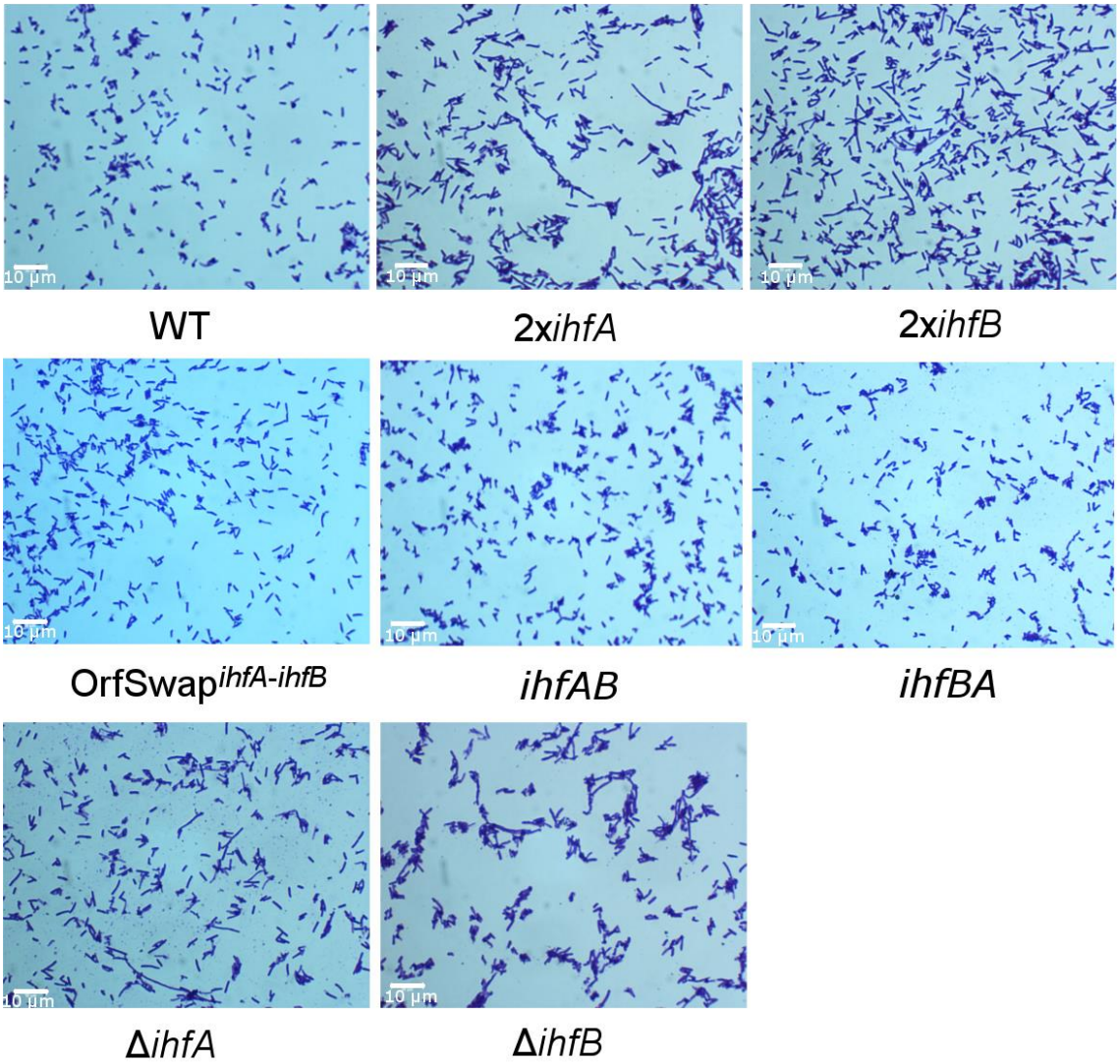


Figure 3.4 Morphology of the strains used in the IHF work at mid-exponential growth phase. Bacteria were harvested, washed with PBS, heat-fixed, stained with crystal violet and viewed under 1000x magnification with oil immersion lens. Filamentous cells were observed in the 2x*ihfA*, 2x*ihfB*, Δ*ihfA*, Δ*ihfB* strains. All images are representative of three biological replicates. 10 µm bar is given for reference.

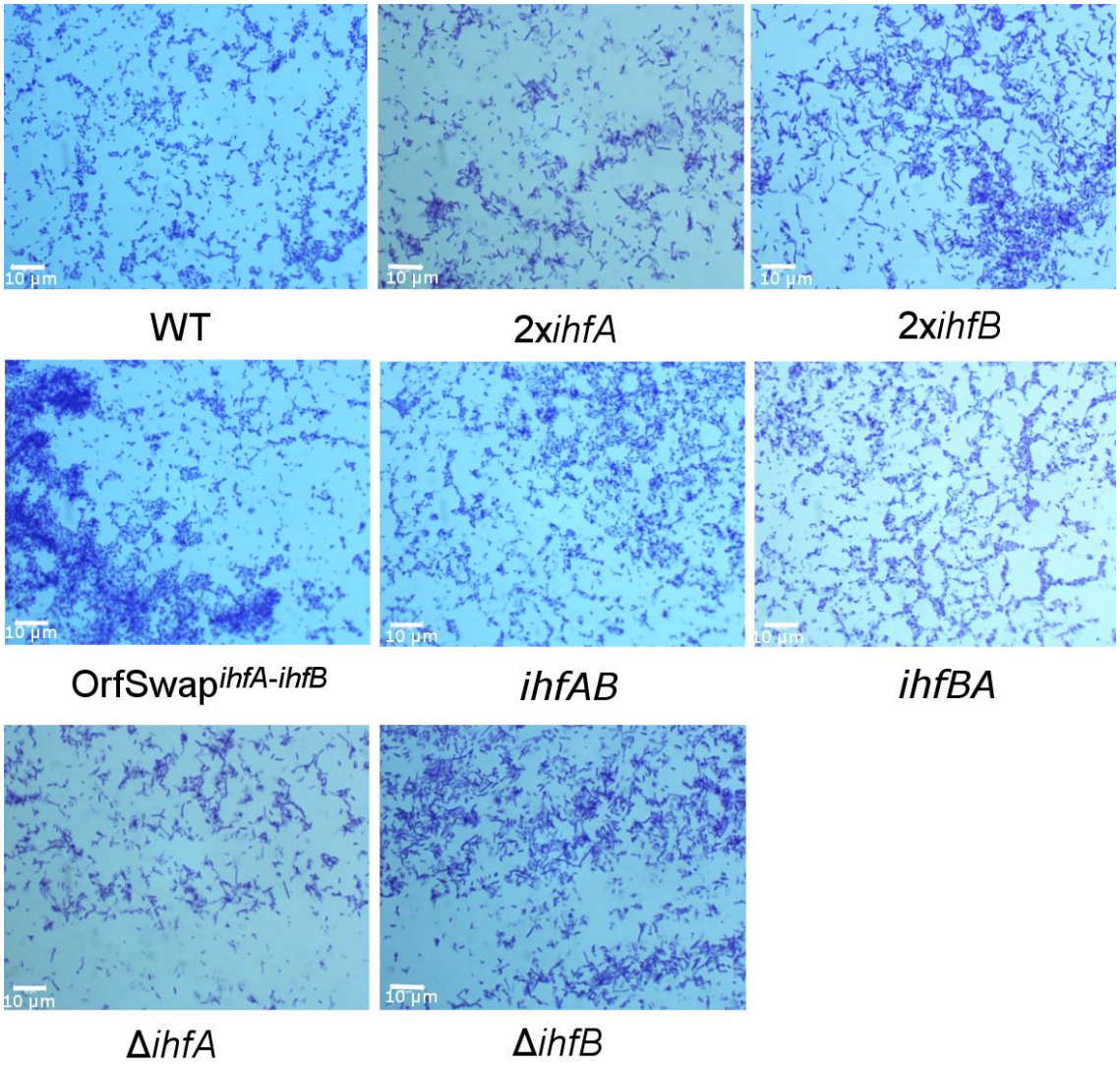


Figure 3.5 Morphology of the strains used in the IHF work at late stationary growth phase. Bacteria were harvested, washed with PBS, heat-fixed, stained with crystal violet and viewed under 1000x magnification with oil immersion lens. Filamentous cells were observed in the 2x*ihfA*, 2x*ihfB*, Δ*ihfA*, Δ*ihfB* strains. All images are representative of three biological replicates. 10 µm bar is given for reference.

3.2.5 Competitive fitness alterations in strains with repositioned *ihf* genes

Measuring growth of a particular strain on its own provides an essential piece of information, but it may not always detect a growth defect. When different strains are co-cultured in the same flask, competition for space and resources occurs and any subtle defects, that were otherwise hidden, manifest themselves. As a result, one of strains could become more or less competitive (Gerganova *et al.*, 2015). The difference in fitness could be further resolved by varying competition growth conditions (Fitzgerald *et al.*, 2015).

A reciprocal ORF exchange of the nucleoid-associated-protein-encoding genes *hns* and *stpA* was made previously. This rewired strain unexpectedly displayed higher competitive fitness than the WT at 37°C and high medium osmolarity (Fitzgerald *et al.*, 2015). In the interests of synthetic biology, this study showed that to create a bacterium with superior properties a single reposition of ORF of important regulatory gene was sufficient. The underlying reason for the fitness increase was premature expression of the RpoS stress-and-stationary-phase regulon.

Competitive fitness of the OrfSwap$^{ihfA\text{-}ihfB}$, *ihfAB* and *ihfBA* strains was determined by co-culturing these strains with the WT in the same flask, beginning with inocula of equal numbers of cells and enumerating each strain after 24 h of growth. To distinguish between the strains, derivatives of each strain were constructed that harboured a chloramphenicol resistance cassette within the non-expressing pseudogene *SL1483* (Kröger *et al.*, 2013) (Table 2.1). The Cm^R cassette enabled selection and did not affect the competitive fitness of the strain that carried it. Competitions were performed with a competitor strain carrying the Cm^R and with the WT carrying the Cm^R as a control, to ensure that any fitness difference observed was due to chromosomal repositionings but was not caused by carriage of the Cm^R within *SL1483*. There was no difference in competitive fitness observed between the WT and any of the IHF repositioned strains (Fig. 3.6a). However, when medium with 0.3 mM NaCl (to introduce high osmolarity) was used, *ihfAB* was outcompeted by the WT in 24 h (Fig. 3.6b). These competitions uncovered a difference in fitness

between the two IHF operons that is manifested upon increase in osmotic pressure due to increased NaCl concentration.

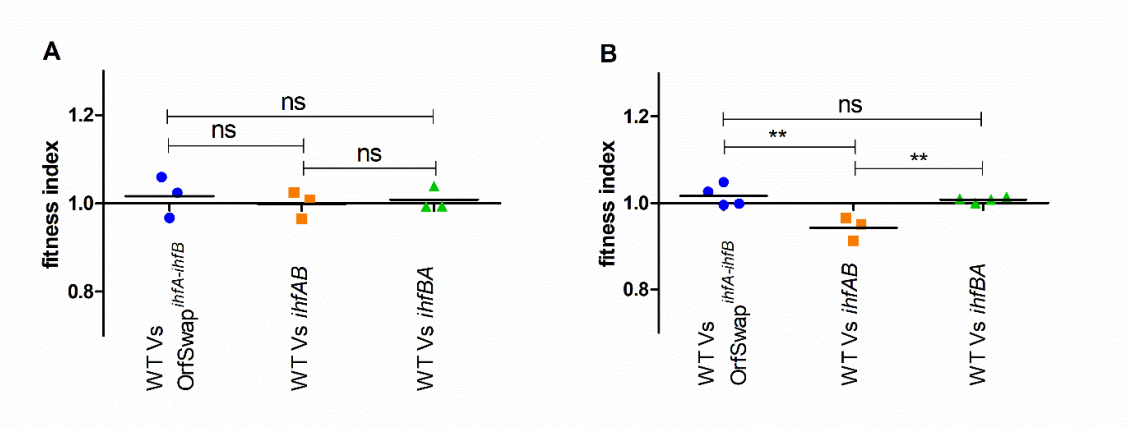


Figure 3.6 Competitive fitness of the strains with repositionings of the *ihf* genes. Fitness of the OrfSwap$^{ihfA\text{-}ihfB}$, *ihfAB* and *ihfBA* strains relative to the WT SL1344 in LB broth supplemented with a) 0.171 mM NaCl or b) 0.3 mM NaCl and grown for 24 h at 200 rpm at 37°C. Fitness index = 1 means that the competed strains were equally fit, f.i. < 1 indicates that the competitor strain (OrfSwap$^{ihfA\text{-}ihfB}$, *ihfAB* or *ihfBA*) was less fit than the WT, f.i. > 1 indicates that the competitor was more fit than the WT. The *ihfAB* operon strain was less fit than the WT at 0.3 mM NaCl. One-way ANOVA with Tukey's Multiple Comparison Test was used to determine significance, where $p<0.05$.

3.2.6 Effect of *ihf* gene repositionings on cell motility

Motility in *Salmonella* depends upon production and activity of flagella (Fitzgerald *et al.*, 2014). As a master regulator, IHF influences the expression of multiple genes, including that of the entire motility regulon: IHF is a positive regulator of flagellar gene expression at all stages of flagellum production (Mangan *et al.*, 2006). The motilities of the OrfSwap$^{ihfA\text{-}ihfB}$, *ihfAB* and *ihfBA* strains were compared with the WT by measuring the diameters of motility zones after 5 h of incubation at 37°C on 0.3% LB agar (Fig. 3.7). Motility differences were reported as the fold change relative to the WT.

In line with the known positive role of IHF in motility, the OrfSwap$^{ihfA\text{-}ihfB}$ displayed a statistically significant decrease in motility, however, the *ihfAB* and the *ihfBA* displayed motility comparable to that of the WT (Fig. 3.7). This indicates that in the strains where the subunits of IHF heterodimer are produced in a transcriptionally coordinated fashion from the same operon, motility was unaffected. Whereas in the OrfSwap$^{ihfA\text{-}ihfB}$ strain where the subunits of the heterodimer were produced under the control of each other's regulatory inputs, IHF fails to regulate the genes involved in motility as efficiently as in the WT. The motility phenotype of the OrfSwap$^{ihfA\text{-}ihfB}$ strain resembled that of *ihfA ihfB* double mutant (Mangan *et al.*, 2006).

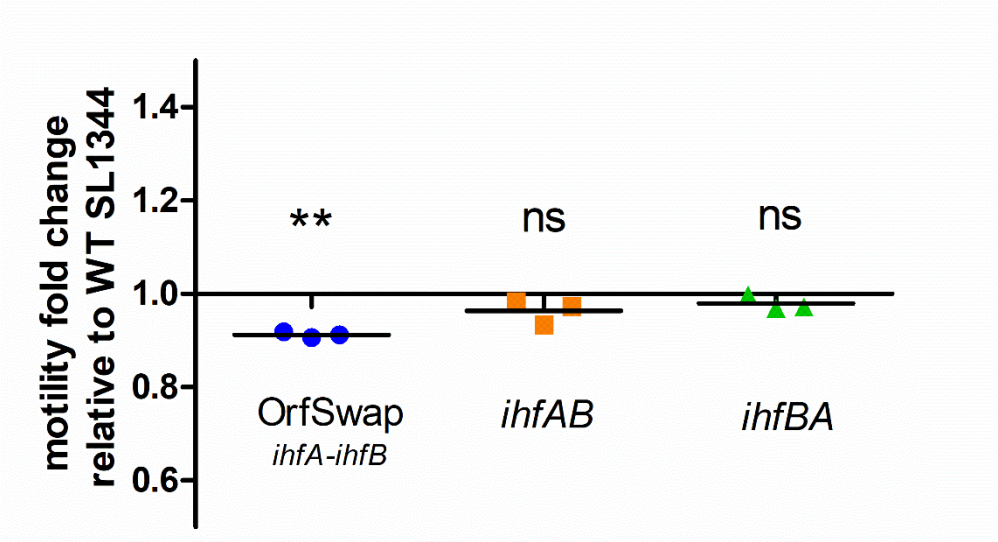


Figure 3.7 Motility of the strains with repositionings of the *ihf* genes relative to the WT. Diameters of swimming motility zones were measured after 5 h incubation at 37°C on soft 0.3% LB agar. The OrfSwap$^{ihfA\text{-}ihfB}$ strain was less motile than the WT, while the *ihfAB* and the *ihfBA* strains were equally motile. Values below 1 indicate that the strain was less motile than the WT. Significance was found by one sample T-test, where $P<0.05$.

3.2.7 Effect of *ihf* gene repositionings on SPI-1 and SPI-2 gene expression

Salmonella pathogenicity islands (SPIs) are long, horizontally acquired genetic regions on the *Salmonella* chromosome (Schmidt & Hensel, 2004). SPI-1 and SPI-2 are the most studied SPIs that each encode a type III secretion system. These mediate *Salmonella's* virulence via invasion of epithelial cells and survival within macrophages respectively (Galán, 1996; Shea *et al.*, 1996). Expression of both SPIs is regulated by complex regulatory networks that are interconnected (Fass & Groisman, 2009; Cameron & Dorman, 2012). IHF is one of the global regulators that positively affects expression of both SPI-1 and SPI-2. IHF levels during growth cycle may coordinate SPI-1 and SPI-2 expression (Mangan *et al.*, 2006).

To evaluate an effect of the chromosomal repositionings of *ihfA* and *ihfB* on the expression of SPI-1 and of SPI-2, *gfp* chromosomal fusions to the promoters of *prgH* and of *ssaG*, genes that encode needle components of SPI-1 and SPI-2, respectively, were used (Ibarra *et al.*, 2010). Expression of these genes is archetypal as it reflects expression patterns of the entire corresponding SPI (Kröger *et al.*, 2013). The fusions were transduced into the WT, OrfSwap$^{ihfA\text{-}ihfB}$, *ihfAB, ihfBA*, Δ*ihfA* and Δ*ihfB* strains by P22-mediated generalized transduction, as in 2.10 (Table 2.1).

GFP fluorescence at 528 nm was measured every 20 min over 24 h and divided by OD_{600} for normalization. The behaviour of the reporter fusions in the WT was in agreement with published data (Sridhar & Steele-Mortimer, 2016) – WT SPI-1 expression peaked during mid-exponential growth phase, while SPI-2 expression peaked from early stationary phase and plateaued thereafter (Fig. 3.8). Substantial decreases in both SPI-1 and SPI-2 expression were seen in strains with the individual IHF subunit gene deletions, confirming that IHF is needed for the activation of both SPI-1 and SPI-2 gene expression (Fass & Groisman, 2009). These data indicate that only heterodimeric IHF can successfully activate both SPIs. In the OrfSwap$^{ihfA\text{-}ihfB}$ strain, expression of both pathogenicity islands was almost identical to that in the WT except for a small, statistically significant, decrease during late-stationary phase in SPI-1 (Fig.

3.8a) and during mid-exponential phase in SPI-2 (Fig. 3.8c). SPI-1 expression differed between the two *ihf* operon strains – *ihfAB* expressed SPI-1 in the same way as the WT, while *ihfBA* expressed SPI-1 at a significantly lower level during exponential phase; there was also a small, statistically significant. decrease during late-stationary phase (Fig. 3.8b). SPI-2 expression in the *ihfAB* and *ihfBA* operon strains was identical to that of the WT (Fig. 3.8d).

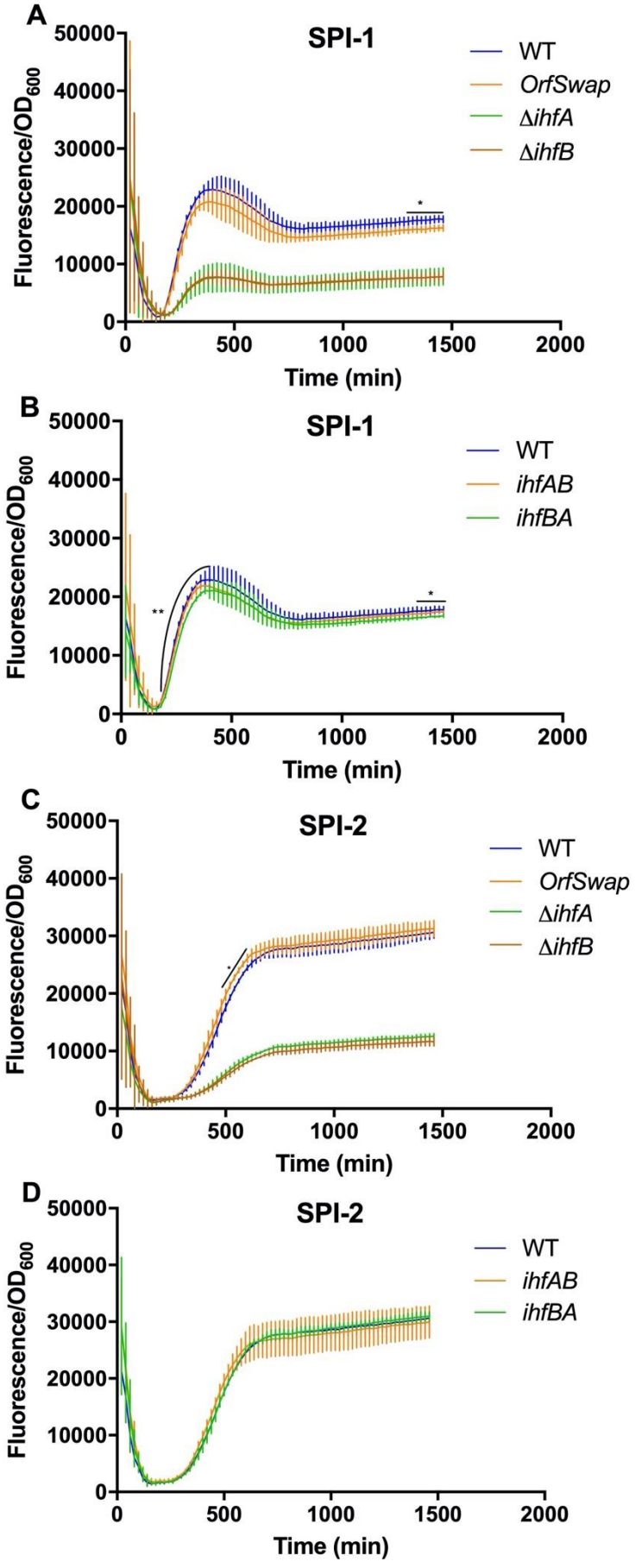

A
SPI-1
WT
OrfSwap
ΔihfA
ΔihfB
Fluorescence/OD600
Time (min)
B
SPI-1
WT
ihfAB
ihfBA
Fluorescence/OD600
Time (min)
C
SPI-2
WT
OrfSwap
ΔihfA
ΔihfB
Fluorescence/OD600
Time (min)
D
SPI-2
WT
ihfAB
ihfBA
Fluorescence/OD600
Time (min)

Figure 3.8 *Salmonella* pathogenicity islands expression measured with aid of *gfp* reporter fusions in strains with *ihf* gene repositionings compared to the WT. Fluorescence at 528 nm normalized against OD_{600} was plotted for every 20 min measurement over 24 h. a) SPI-1 expression in the OrfSwap$^{ihfA\text{-}ihfB}$ strain was slightly lower than in the WT during the late stationary growth phase. SPI-1 expression was lower in Δ*ihfA* and Δ*ihfB*. b) SPI-1 expression in the *ihfAB* was identical to the WT. SPI-1 expression in the *ihfBA* was lower than the WT at the exponential growth phase and at the late stationary growth phase. c) SPI-2 expression in the OrfSwap$^{ihfA\text{-}ihfB}$ strain was slightly lower than in the WT during the mid-exponential phase. SPI-2 expression was lower in the Δ*ihfA* and Δ*ihfB* strains. d) SPI-2 expression in the *ihfAB* and *ihfBA* operon strains was identical to the WT. All plots are results of at least three biological replicates, error bars represent standard deviation. Asterisks represent significance for timepoints highlighted with black curves. Significance was found by Student's unpaired T-test, where $P<0.05$.

3.2.8 Ability of the *Salmonella* strains with *ihf* gene repositionings to infect and survive in murine macrophages

The *Salmonella*-containing vacuole of the mammalian macrophage is an important niche for *Salmonella* during infection (Finlay & Brumell, 2000). So, while the results of various *in vitro* phenotypic tests are useful, it is important to understand that no matter how well and seemingly infection-relevant condition is chosen for *in vitro* phenotypic tests, it will still remain artificial – one that the bacterium never encounters in its natural environment. Thus, it was decided to test the properties of the rewired strains under the influence of a mixture of hostile responses *in vivo*, including the production of hydrogen peroxide, other reactive oxygen and nitrogen species, acidic pH, antimicrobial peptides, etc., to which bacteria are exposed when inside a macrophage.

An important factor that enables *Salmonella* infection is SPI gene expression (Galán, 1996). SPI-1 is activated in the gut lumen in response to low oxygen and high osmolarity and exploited for entry into epithelial cells and macrophages (Bajaj *et al.*, 1996). SPI-2 is activated when *Salmonella* is inside a macrophage vacuole in response to acidic pH, low Mg^{2+} and Fe^{2+} to ensure bacterial survival, replication and eventually, dissemination (Lee *et al.*, 2000; Quinn *et al.*, 2014; Fass & Groisman, 2009).

There was a small difference in SPI-1 and SPI-2 expression at some growth phases in the strains where *ihfA* and *ihfB* are repositioned (Fig. 3.8), indicating that despite IHF's participation in SPI gene regulation, their expression was not greatly affected. However, the small change observed made it feasible to carry out assays to test the performance of these strains in live infection models.

RAW264.7 murine macrophages were chosen as they constitute a well-studied cell line that is frequently used for studying *Salmonella* infection. Gentamycin protection assays, as in 2.22, were carried out to assess the performance of the bacterium at different stages of the infection process: efficiency of entry to the macrophage, intracellular survival and replication over a 20-h period post-infection. Survival and replication of *Salmonella* in macrophages is dependent upon the mechanism of entry (Drecktrah *et al.*, 2006). Two main models of entry utilised by *Salmonella* are: (1) – SPI-1 mediated active invasion and (2) –

passive host-mediated phagocytosis. Both models are important as they occur in *Salmonella'*s natural environment. Macrophages can encounter highly invasive *Salmonella*, in which SPI-1 is induced, when they first meet as the bacteria emerge from the intestinal epithelial cell barrier (Finlay & Brumell, 2000; Knodler *et al.*, 2010) (Fig. 1.7). But *Salmonella* also can be taken up by a macrophage via phagocytosis. This occurs when it escapes its original macrophage replication niche while expressing the SPI-2 type III secretion system needle on its surface or when *Salmonella* is spread to the circulatory system during systemic infection (Cirillo *et al.*, 1998).

To achieve active entry to macrophage via SPI-1, bacteria were grown to mid-exponential growth phase in LB broth to maximize SPI-1 expression (Steele-Mortimer *et al.*, 1999). To ensure passive entry via phagocytosis, bacteria were grown to late stationary growth phase and complement-opsonised (Lee & Falkow, 1990). Invasive SPI-1 induced bacteria were used at MOI 5, while opsonised SPI-2 induced bacteria were used at MOI 20 to achieve similar entry numbers. The 1-h post-infection timepoint was chosen as a reference point for invasion.

The infection model with SPI-2-induced and opsonised *Salmonella* to promote phagocytosis was used to infect RAW264.7 macrophages. Both infection entry efficiency and long-term intracellular survival were tested. Colony Forming Units (CFUs) inside macrophages were counted at 1 h and normalized against CFUs in the infection mix followed by normalisation against the WT. The data obtained showed that the OrfSwap$^{ihfA\text{-}ihfB}$ strain was phagocytosed more readily than the WT (Fig. 3.9a). This enhanced phagocytosis suggested that the cell membrane composition of the OrfSwap$^{ihfA\text{-}ihfB}$ is different from that of the WT. This may have allowed more efficient complement deposition on the microbial cell surface and hence better phagocytosis.

Salmonella survival and replication inside a macrophage was tested using the same infection model. CFUs inside macrophages were counted at further timepoints and normalized against CFUs at 1 h. For all strains, after slight initial growth at 3 h post infection, a killing event was observed at 8 h p.i. Later, at 16 h p.i., *Salmonella* appeared to adapt and replicate significantly, followed by a

further decline by 20 h p.i., possibly caused by macrophage cell death (van der Velden *et al.*, 2000). When the repositioned strains were compared to the WT, no significant differences were found, at any of the timepoints tested (Fig 3.9b).

The lack of any difference in survival rates between the strains may seem to be inconsistent with the observation that the expression levels of the SPI genes in the repositioned strains were slightly lower than in the WT (Fig. 3.8). However, SPI expression depends on the mechanism of *Salmonella* uptake by the macrophage. Following phagocytosis of *Salmonella*, both SPI-1 and SPI-2 regulon genes are expressed to a much smaller magnitude than when *Salmonella* invades macrophages using the SPI-1-encoded T3SS (Drecktrah *et al.*, 2006). Therefore, an infection model with SPI-1 induced bacteria was tested. No differences in the active SPI-1 mediated entry was found between the WT and the strains with the repositioned IHF-encoding genes (Fig. 3.10). This suggests that the magnitude of the SPI expression changes observed in Fig. 3.8 was too low to cause detectable alterations in the ability to infect mammalian macrophages.

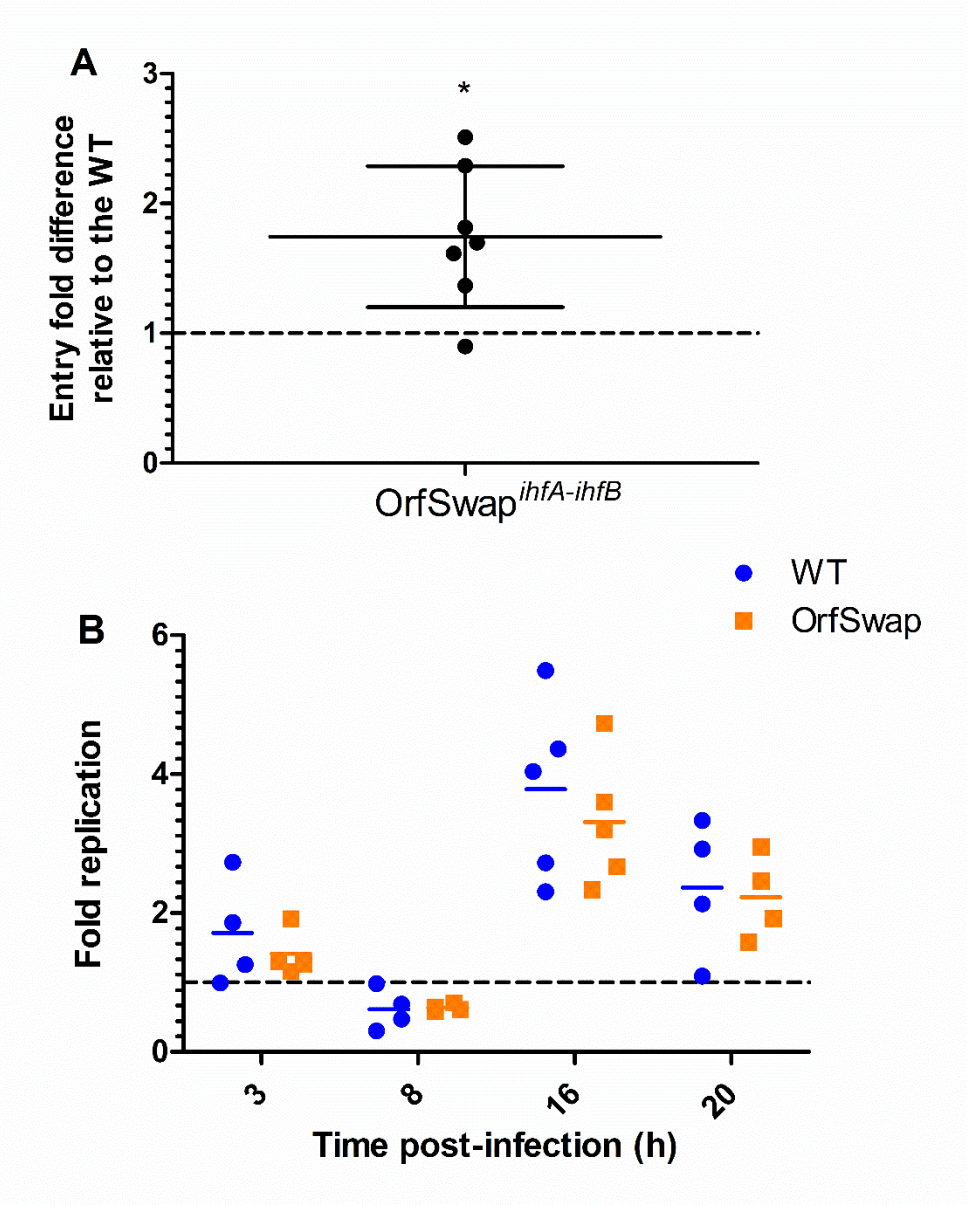


Figure 3.9 Entry by phagocytosis and survival of *Salmonella* strains OrfSwap$^{ihfA\text{-}ihfB}$ *ihfAB* and *ihfBA* in RAW264.7 cells. Cells were infected with SPI-2 induced bacteria, grown to the stationary phase and complement-opsonised to promote phagocytosis. a) Entry was measured by enumerating CFUs 1 h post-infection. Each replicate was normalized to the infection mix and the WT. Mean and standard deviation are shown. Significance was found by one sample T-test, where $P<0.05$. b) Survival and replication were measured by enumerating CFUs at 3 h, 8 h, 16 h and 20 h post-infection. Fold replication represents the CFU recovered at the particular time point divided by the CFU at 1 h. Mean and individual replicates are shown. No significant difference was found by unpaired Student's T-test.

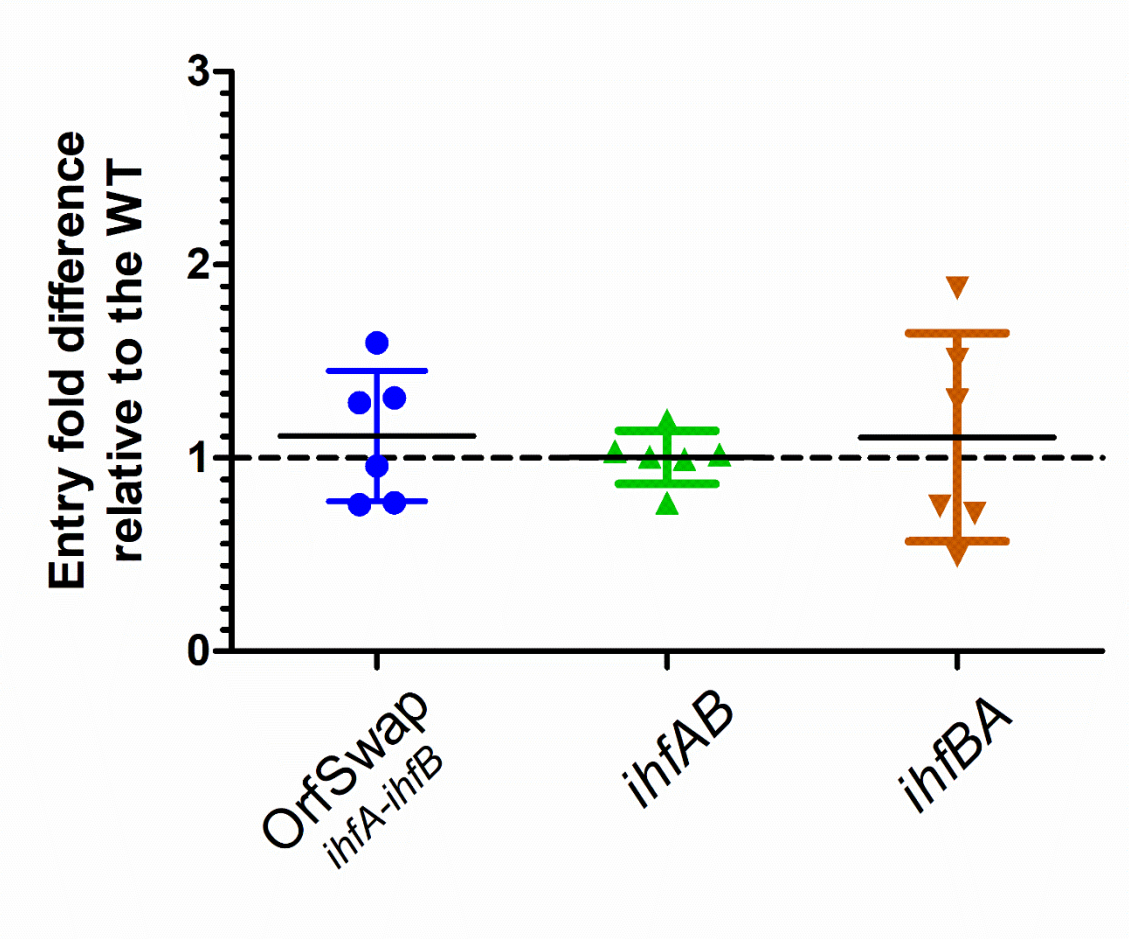


Figure 3.10 SPI-1 mediated entry of *Salmonella* strains OrfSwap$^{ihfA\text{-}ihfB}$, *ihfAB* and *ihfBA* in RAW264.7 cells. Cells were infected with SPI-1 induced bacteria, grown to the mid-exponential phase of growth to promote SPI-1 mediated active entry. a) Entry was measured by enumerating CFUs 1 h post-infection. Each replicate was normalized to the infection mix and the WT. Mean and standard deviation are shown. No significant differences were found by one sample T-test, where $P<0.05$.

3.3 Discussion

In order to learn about the significance of chromosomal positions of *ihfA* and *ihfB*, and about the significance of their expression as stand-alone genes rather than as an *ihf* operon, the open reading frames (ORFs) of these two genes were repositioned in different ways. The first kind of reposition followed the strategy used previously for studying gene positions (Fitzgerald *et al.*, 2015). The ORFs of *ihfA* and *ihfB* were reciprocally exchanged to create the OrfSwap$^{ihfA\text{-}ihfB}$ strain in which *ihfA* expression was controlled by the regulatory regions of *ihfB*, including its promoter, its upstream regulatory sites and its transcriptional terminator; the analogous reciprocal exchange was made for the *ihfB* ORF. The second kind of repositioning was novel and involved creation of the *ihfAB* and *ihfBA* artificial operons. These arrangements were chosen to test the effect of seemingly advantageous co-expression of subunits of the same heterodimeric protein. To create *ihfBA*, the ORF of *ihfA* including a 20 b.p. region upstream, that contained a ribosome binding site, was inserted downstream of *ihfB*. The original *ihfA* gene was deleted to avoid having two copies of *ihfA* in the same organism. *IhfAB* was created by an analogous set of manipulations.

The entire *fis* and *dps* genes, including their regulatory regions, have been exchanged in *Salmonella* (Bogue *et al.*, 2020) to reveal that when the entire gene is moved to a new location, fewer changes in *Salmonella's* phenotypes, gene expression and global FIS binding take place, unlike when the ORFs alone are exchanged. It turned out to be impossible to compare the effects of ORF and gene exchanges of *ihfA* and *ihfB* because each of these genes is embedded in a set of complex regulatory circuits with overlapping transcription units and promoters (Weglenska *et al.*, 1996; Wertheimer *et al.*, 1988). The regulatory regions of *ihfA* and *ihfB* overlap other genes in their respective heterologous operons, so a whole gene exchange in any configuration would have affected neighbouring genes with unpredictable effects on cell physiology.

The experiments described in this chapter dealt exclusively with the construction and the phenotypic consequences of IHF-encoding gene repositionings. All the strains with IHF repositionings were constructed using a

previously published method (Datsenko & Wanner, 2000). Whole genome sequencing was performed on each strain and revealed no compensatory mutations.

IHF is essential for site-specific recombination and exerts its effect via sharp DNA bending. This NAP was discovered as a factor required for integration of phage lambda into the genome of *E. coli* by site-specific recombination (Miller & Friedman, 1980; Dorman & Bogue, 2016). Flp-mediated (flippase) site-specific recombination was used throughout the work for removal of antibiotic resistance casettes flanked by Flp recognition targets (FRT) sequences. This method utilized recombination between the FRT sites catalysed by the flippase (Cherepanov & Wackernagel, 1995). Individual *ihf* deletions and the 2x intermediate strains confirmed the role of IHF in site-specific recombination as it was impossible to use this method for the removal of antibiotic resistance cassettes from the 2x*ihfA*, 2x*ihfB*, *ΔihfA* and *ΔihfB* strains.

IHF performs gene regulatory and architectural functions through its binding to DNA at consensus binding sites (Section 1.3) and altering DNA path by sharply bending it (Fig. 1.3) (Goodrich *et al.*, 1990; Rice *et al.*, 1996). One of important DNA architectural roles is a positive effect that IHF plays during intiation of replcation at the origin of chromosomal replication. Binding and bending of DNA at the origin by IHF promotes binding of DnaA that in turn oligomerises and unwinds AT-rich DUE (DNA unwinding element) that leads to a series of events that conclude at replisome formation (Ryan *et al.*, 2004). The lack of IHF or expression of not fully-functional IHF may thus lead to an adverse effect on the growth cycle and the cell cycle. No effect of individual subunit deletions of IHF on bacterial growth was shown before (Mangan *et al.*, 2006; Prieto *et al.*, 2012) but growth in those studies was measured by OD_{600}. In the present study, there was similarly no difference in growth between the WT and the strains with repositioned or deleted *ihfA* and *ihfB* when measured by absorbance at 600_{nm} (Fig. 3.2). However, viability count measurements displayed a growth defect in strains with individual *ihf* deletions (Fig. 3.3). This result highlights the importance of fully functional IHF for the growth cycle. At the same time, there were no growth changes observed in the OrfSwap$^{ihfA\text{-}ihfB}$, *ihfAB* or *ihfBA* strains. Problems with the cell cycle are likely to manifest as morphological aberrations.

Indeed, filamentous cells were visible in the strains where either of the IHF subunits was absent at mid-exponential and late stationary phases, while the strains with the repositioned IHF-encoding genes displayed no morphological changes (Fig. 3.4 and Fig. 3.5). Lack of changes in growth and morphology suggests that in the strains with repositioned *ihfA* and *ihfB,* IHF is produced in a form that is able to function correctly during the initiation of chromosomal replication.

In previous studies, strains in which the positions of different NAP-encoding genes were altered displayed positive and negative alterations in competitive fitness previously (Gerganova *et al.*, 2015; Fitzgerald *et al.*, 2015). In the strains with repositioned *ihfA* and *ihfB*, no competitive fitness differences were found under the standard growth conditions (Fig. 3.6a). However, in the medium with elevated NaCl concentration *ihfAB* was less fit than the WT (Fig. 3.6b). This shows that there is a hidden difference between the IHF operon strains that is manifested in an environment with high osmolarity.

It was determined previously that IHF positively regulates flagellar motility in *Salmonella*. Single and double IHF subunit deletions result in transcriptional changes in multiple motility-asociated genes, and the reduction in motility was observed to correlate with downregulation of the genes involved in the final stages of flagellum production (Mangan *et al.*, 2006). In the present study, motility measurements carried out in the strains with the repositioned IHF-encoding genes showed that it was decreased in the OrfSwap$^{ihfA\text{-}ihfB}$ but not in the *ihfAB* or *ihfBA* strains (Fig. 3.7). This phenotype suggests that there may be changes in expression of genes associated with motility in OrfSwap$^{ihfA\text{-}ihfB}$. Motility is also the first IHF-linked phenotype that displayed a difference between the types of rearrangements as no motility defect was detected in the strains with artificial *ihf* operons.

One of the most prominent features that distinguish the genome of *Salmonella* from the genome of *E. coli* is the presence of *Salmonella* pathogenicity islands. These horizontally acquired AT-rich regions are responsible for the entry into epithelial cells and macrophages (SPI-1) and for the survival in, and the escape from macrophages (SPI-2) (O'Croinin *et al.*, 2006). IHF, along with the other

NAP FIS, is known to positively regulate the expression of genes in both pathogenicity islands (Fass & Groisman, 2009). Global expression analysis by microarrays showed that in the absence of IHF, SPI-1 and SPI-2 gene expresion was greatly decreased during early to late exponential growth phase, when IHF is expressed at its highest levels (Mangan *et al.*, 2006). Interestingly, the same study showed that deletions of individual IHF subunits resulted in different gene expression changes, often less significant than the double deletion. For example, there was no change in SPI-2 expression in the Δ*ihfB*, but in the Δ*ihfA* SPI-2 expression was equivalent to that seen in the full absence of IHF. When, in the present study, the expression of archetypal SPI-1 and SPI-2 genes was monitored in the strains with IHF repositions, very little change was found (Fig. 3.8). This is in contrast to the situation in the Δ*ihfA* and the Δ*ihfB* where significant and equivalent decreases in both SPI-1 and SPI-2 gene expression were seen. The most significant (although small) difference found in the repositioned strains was a decrease in SPI-1 expression found in the *ihfBA*; in contrast, SPI-1 expression in the *ihfAB* resembled that of the WT (Fig. 3.8b). Although these results indicate that alterations of the positions of individual IHF-encoding genes do not generally affect the ability of IHF to regulate SPI-1 and SPI-2 expression, they reveal changes to SPI gene expression in the two strains with synthetic *ihf* operons.

It should be noticed that the level of SPI expression is not likely to be uniform for all of the cells in a population. Instead, each individual cell is expected to display a separate expression program. Such expression heterogeneity is known to be true, in particular, for SPI-1 (Clark *et al.*, 2011; Hautefort *et al.*, 2003). GFP fluorescence measurements from the reporter fusions display only the averaged fluorescence values from the entire population, therefore, slight variations in SPI expression that were deemed statistically significant (Fig. 3.8) might be not significant biologically. It would be more valuable to carry out flow cytometry to provide an answer whether the proportion of SPI-1 ON/OFF cells differs between the WT and the strains with the repositioned genes.

The effect of *ihf* gene repositionings on *Salmonella*'s most unique phenotype – the ability to infect and survive within a macrophage – was tested. Two modes of entry were studied – SPI-1 induced active entry and host-mediated passive

entry by phagocytosis (Drecktrah *et al.*, 2006). The OrfSwap$^{ihfA\text{-}ihfB}$ strain was found to be phagocytosed more efficiently than the WT, albeit with no difference in long-term survival and replication beween the OrfSwap$^{ihfA\text{-}ihfB}$ and the WT. The efficiency of phagocytosis is determined by the envelope composition of the bacterium. *Salmonella* surface antigens are recognised by complement, promoting more efficient recognition and uptake by macrophages. Changes in structure of these antigens, such as the O-antigen of lipopolysaccharide (LPS), affect the efficiency of phagocytosis (Murray *et al.*, 2006). Enhanced phagocytosis of the OrfSwap$^{ihfA\text{-}ihfB}$ strain may be a result of changes in its outer membrane composition. There were no discernible differences in SPI-1 mediated active entry between the WT and any of the strains with the altered positions of the genes that encode IHF. This result is supported by the little to no difference in SPI-1 and SPI-2 expression between the WT and these strains.

In contrast to this study, genes that encode different NAPs were in focus of the previous studies where genes were reciprocally exchanged to find out more about significance of gene order and test resilience of bacteria to such repositions. For example, the orthologous NAPs H-NS and StpA were exchanged in (Fitzgerald *et al.*, 2015) and the heterologous NAPs FIS and Dps were exchanged in (Bogue *et al.*, 2020). Both studies concluded that those exchanges are tolerated well by *Salmonella*. In the present study, genes encoding the subunits of the same heterodimeric NAP (IHF) were rearranged. Although IHF plays multiple important roles in *Salmonella*, reciprocal rearrangements of *ihfA* and *ihfB* did not cause many significant changes to bacterial wellbeing. This is possibly because although *ihfA* and *ihfB* are geographically dispersed on the chromosome, their regulation is very similar – both genes are negatively controlled by ppGpp and negativelly autoregulated (Lemke *et al.*, 2011; Aviv *et al.*, 1994). Therefore *ihfA* transcription can be equally efficient when initiated from its own promoter or from the P_{ihfB}, on its own or in an operon with *ihfB*, and *vice versa*.

The phenotypic tests described in this chapter showed that the repositionings of *ihfA* and *ihfB* were tolerated well by *Salmonella*. Some experiments revealed differences that were associated with the reciprocal ORF exchange strain. The strains with the synthetic *ihf* operons were found not to be equivalent. Gene

expression and protein production studies will be described in Chapter 4 to finalize our understanding of the significance of the IHF-encoding genes positions in *Salmonella*.

Chapter 4:

Repositioning of *ihfA* and *ihfB* in *Salmonella enterica* serovar Typhimurium and its effect on *ihfA* and *ihfB* gene expression and IHF protein production

4.1 Introduction

No additional mutations were found in the genomes of the strains with *ihf* gene repositionings (section 3.2.2). This means that all the phenotypic characteristics described in Chapter 3 are the direct outcomes of the IHF expression alterations that were caused by the repositionings. To finalize the analysis of the significance of the chromosomal positions of *ihfA* and *ihfB* ORFs, the results of the IHF-associated phenotype testing that were presented in the previous chapter were supplemented with gene expression and protein production studies.

In the simplistic model of the ORF exchange, the ORF under a control of the regulatory region of another gene would accept the offered regulatory input to display gene expression level and patterns attributed to the original ORF. This exchange in the gene expression characteristics would in turn result in the exchange in the dynamics of protein production. Such reciprocal exchanges of the gene expression and of the protein production patterns were seen in the strains where the ORFs of NAPs with the contrasting gene expression characteristics were exchanged, for example, *fis* and *dps* ORF exchange or *hupA* and *hupB* ORF exchange (gene expression tested only) (Bogue *et al.*, 2020). When ORFs of the genes whose expression characteristics resembled each other (*hns* and *stpA*) were exchanged, the resulting gene expression and protein production patterns could not be simply interpreted as exchanges but were unique to the created strain (Fitzgerald *et al.*, 2015). Since *ihfA* and *ihfB* encode subunits of one protein, their expression patterns are expected to be highly similar in the WT, however outcomes of the ORF exchange of the genes that encode IHF are difficult to predict. Both the gene expression patterns of the repositioned IHF-encoding genes and the protein production patterns of IhfA and IhfB from the resulting hybrid loci were directly assessed with aid of RT-qPCR and Western blotting. RNA stability assays were carried out to investigate the reasons for the disagreement between the discovered gene expression and protein production patterns.

Testing the global effect of the ORF exchange of the IHF-encoding genes is the other main theme of this chapter. Since IHF is a known regulator of multiple

genes (Swinger & Rice, 2004; Mangan *et al.*, 2006; Prieto *et al.*, 2012), perturbations of the chromosomal positions of the genes that encode it can be expected to affect global gene expression and, as a consequence, the global pattern of protein production. As gene regulation of IHF is mediated by the DNA binding and bending properties of this NAP (Ali Azam *et al.*, 1999; Rice *et al.*, 1996), the ORF exchange of *ihfA* and *ihfB* may also lead to alteration of the global IHF chromosomal binding. Although the potential effect of the repositioning on the global gene expression and IHF-binding patterns is of interest, it was decided to measure global protein production as a proxy for the changes that occur in a cell due to the reciprocal ORF repositioning. Comparing the global gene expression and protein binding in the strain with the chromosomal repositioning to those in the WT are powerful methods for elucidating global effects of the repositioning on a bacterial cell. The proteomic approach was, however, implemented because proteins do most of the work in a cell and because proteins are a final outcome of gene expression. It was also considered to be the most time, labour and cost-effective. This experiment was performed to serve two purposes: first – to provide a picture of the changes in a cell between the WT and the OrfSwap$^{ihfA\text{-}ihfB}$ strains on the proteome level by comparing relative levels of the individual proteins, and second – to confirm the higher abundance of the IhfB subunit over that of the IhfA subunit of IHF at the early stationary phase of growth in both the WT and the OrfSwap$^{ihfA\text{-}ihfB}$ by relative quantification of these subunits within those strains.

No proteome-wide examinations were performed previously in the studies in which genes have been moved to a new location on the bacterial chromosome (Fitzgerald *et al.*, 2015; Gerganova *et al.*, 2015). Nevertheless, the transcriptome of the strain with a reciprocal ORF exchange of *hns* and *stpA* genes was determined and compared to that of the WT to identify only 26 differentially regulated genes (Fitzgerald *et al.*, 2015). The global gene expression was examined in the strain with the *fis* and *dps* gene exchanges (Bogue *et al.*, 2020). It was discovered that the ORF exchange resulted in considerably larger number of upregulated and downregulated genes compared to the whole gene exchange where only a few changes were detected. Since the ORFs of the genes that encode subunits of the same protein were

exchanged in the OrfSwap$^{ihfA\text{-}ihfB}$ strain, it could be hypothesised that the proteome of this strain would harbour only a few differences, compared to that of the WT, that might be associated with motility and cell outer membrane composition, according to the phenotypic changes described in the Chapter 3.

4.2 Results

4.2.1 Gene expression patterns of *ihfA* and *ihfB* in the strains where these genes are repositioned.

Separately positioned *ihfA* and *ihfB* genes in *Salmonella* and *E. coli* are regulated by separate regulatory regions (Aviv *et al.*, 1994; Weglenska *et al.*, 1996). As a result of the genetic repositionings that were performed in this work, the open reading frames of *ihfA* and *ihfB* were removed from their natural loci and placed under the control of non-native regulatory regions. In the OrfSwap$^{ihfA\text{-}ihfB}$ strain, the ORFs of *ihfA* and *ihfB* were brought under the transcriptional control of each other's 5' and 3' regulatory regions (Fig. 2.1). Such repositionings resulted in the production of the hybrid mRNA species consisting of the 5' and 3' UTRs of one gene and the ORF of the other. In the *ihfAB* strain, the *ihfB* ORF was brought downstream of the *ihfA* ORF and under the control of its 5' and 3' UTRs. While in the *ihfBA* strain, in contrast, the *ihfA* ORF was placed downstream of the *ihfB* ORF and under the control of its 5' and 3' UTRs (Fig. 2.1). In the strains with the artificial *ihf* operons, one hybrid bicistronic mRNA was synthesised instead of the two separate mRNA species.

To test if the ORF rearrangements changed the gene expression patterns of *ihfA* and *ihfB* in the OrfSwap$^{ihfA\text{-}ihfB}$, *ihfAB* and *ihfBA* strains, total RNA was extracted from all strains at the various growth stages, as in 2.5.4, DNase-treated and converted to cDNA, as in 2.12. The abundance of the *ihfA* and *ihfB* transcripts was determined by RT-qPCR, as in 2.13, using pairs of ORF-specific primers. SL_ihfA_qPCR_Pf and SL_ihfA_qPCR_Pr for *ihfA* and SL_2ihfB_qPCR_Pf and SL_2ihfB_qPCR_Prev for *ihfB* (Table 2.2). Levels of the transcripts were normalized against *hemX* – a gene whose expression does not change under the growth conditions used here (Kröger *et al.*, 2013). The specificity of the primers was confirmed, as in 2.13.3 (Fig. 4.1).

In the WT, the pattern of the *ihfA* and *ihfB* mRNA expression resembled that of IHF protein production in that it peaked at 7 h, corresponding to the early stationary growth phase (Fig. 4.2) (Ali Azam *et al.*, 1999). In the OrfSwap$^{ihfA\text{-}ihfB}$ strain *ihfA* gene expression dropped at 5 h and 7 h but increased at 2 h, while *ihfB* gene expression matched that in the WT, except a decrease at 7 h. Note

that in the OrfSwap$^{ihfA\text{-}ihfB}$ strain *ihfA* gene expression refers to the levels of the 5'-*ihfB*[UTR]-*ihfA*[ORF]-3' hybrid mRNA and *ihfB* gene expression refers to the levels of the 5'-*ihfA*[UTR]-*ihfB*[ORF]-3' hybrid mRNA, since qPCR primers are ORF-specific.

In the *ihfAB* strain, the gene expression patterns of *ihfA* and *ihfB* displayed the WT trend, but the gene expression of *ihfA* decreased at 5 h, while that of *ihfB* decreased at 7 h (Fig. 4.3a). In the *ihfBA* strain, the gene expression pattern of *ihfA* and *ihfB* similarly followed the general WT trend but the gene expression of *ihfA* decreased at 5 h and at 7 h. (Fig. 4.3b).

These results indicate that the ORFs of *ihfA* and *ihfB* under the control of 5' and 3' regulatory regions of each other did not exchange gene expression characteristics. Instead, they had acquired unique, rescheduled gene expression patterns.

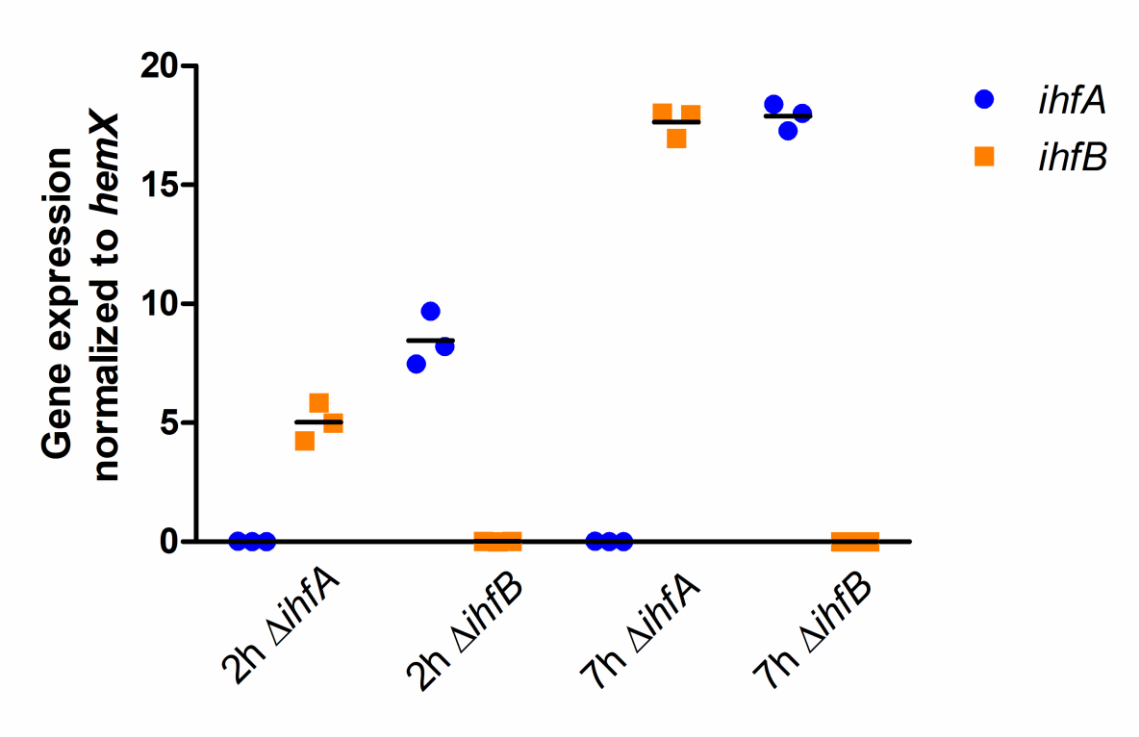


Figure 4.1 Validation of the qPCR primers specific for *ihfA* and *ihfB*. Gene expression of *ihfA* and *ihfB* in the Δ*ihfA* and the Δ*ihfB* strains. No *ihfA* gene expression was detected in Δ*ihfA* and no *ihfB* gene expression was detected in Δ*ihfB* indicating that there is no cross binding of the qPCR primers. The plot is a result of three biological replicates, 2 h and 7 h timepoints represent lag and early stationary growth phases, respectively.

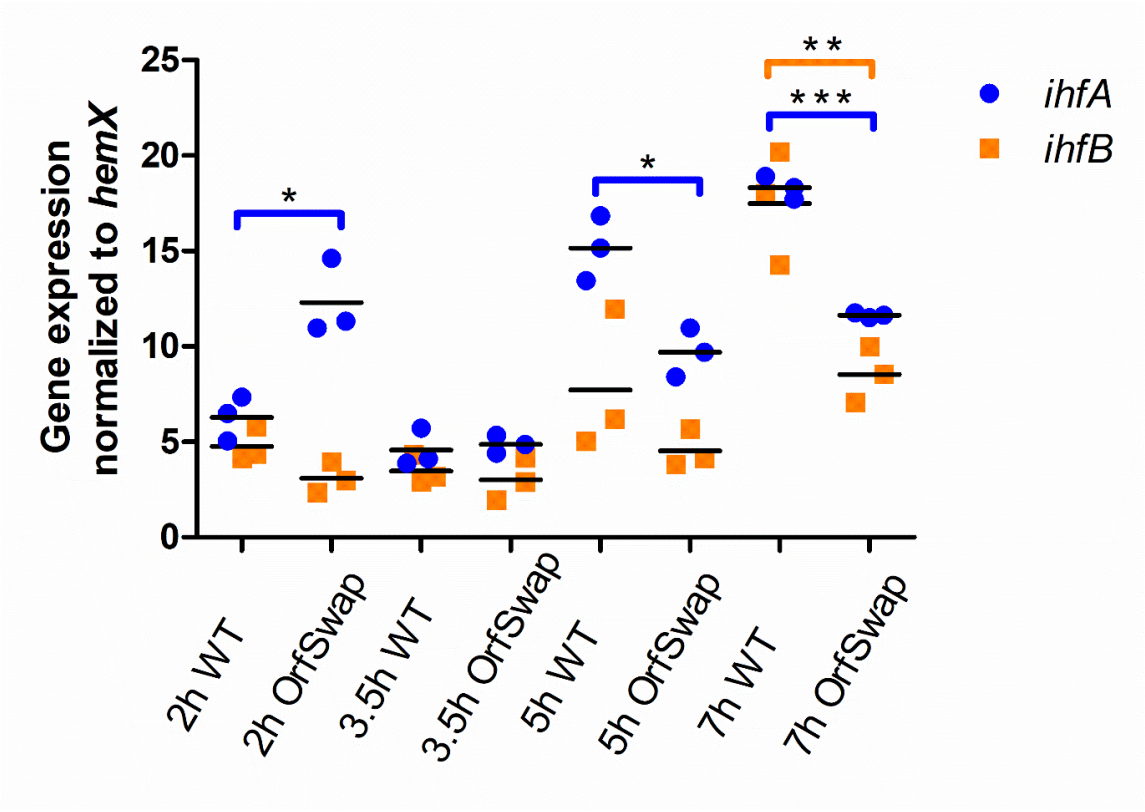


Figure 4.2 Gene expression of *ihfA* and *ihfB* in the SL1344 WT and the OrfSwap$^{ihfA\text{-}ihfB}$ strains. Cells were grown in LB broth at 37°C at 200 rpm and samples were taken at 2 h, 3.5 h, 5 h and 7 h representing lag, exponential, exponential-stationary transition and early stationary phases of growth. *IhfA* gene expression peaked at 5 h and 7 h in the WT, resembling IHF expression, while it was reversed in the OrfSwap$^{ihfA\text{-}ihfB}$ peaking at 2 h. *IhfB* gene expression had comparable levels to that of *ihfA* and increased as bacteria grew. *IhfB* was transcribed to a higher level in the WT than in the OrfSwap$^{ihfA\text{-}ihfB}$ at 5 h and 7 h. All plots are results of three biological replicates, significance was found by Student's unpaired T-test, where $P<0.05$.

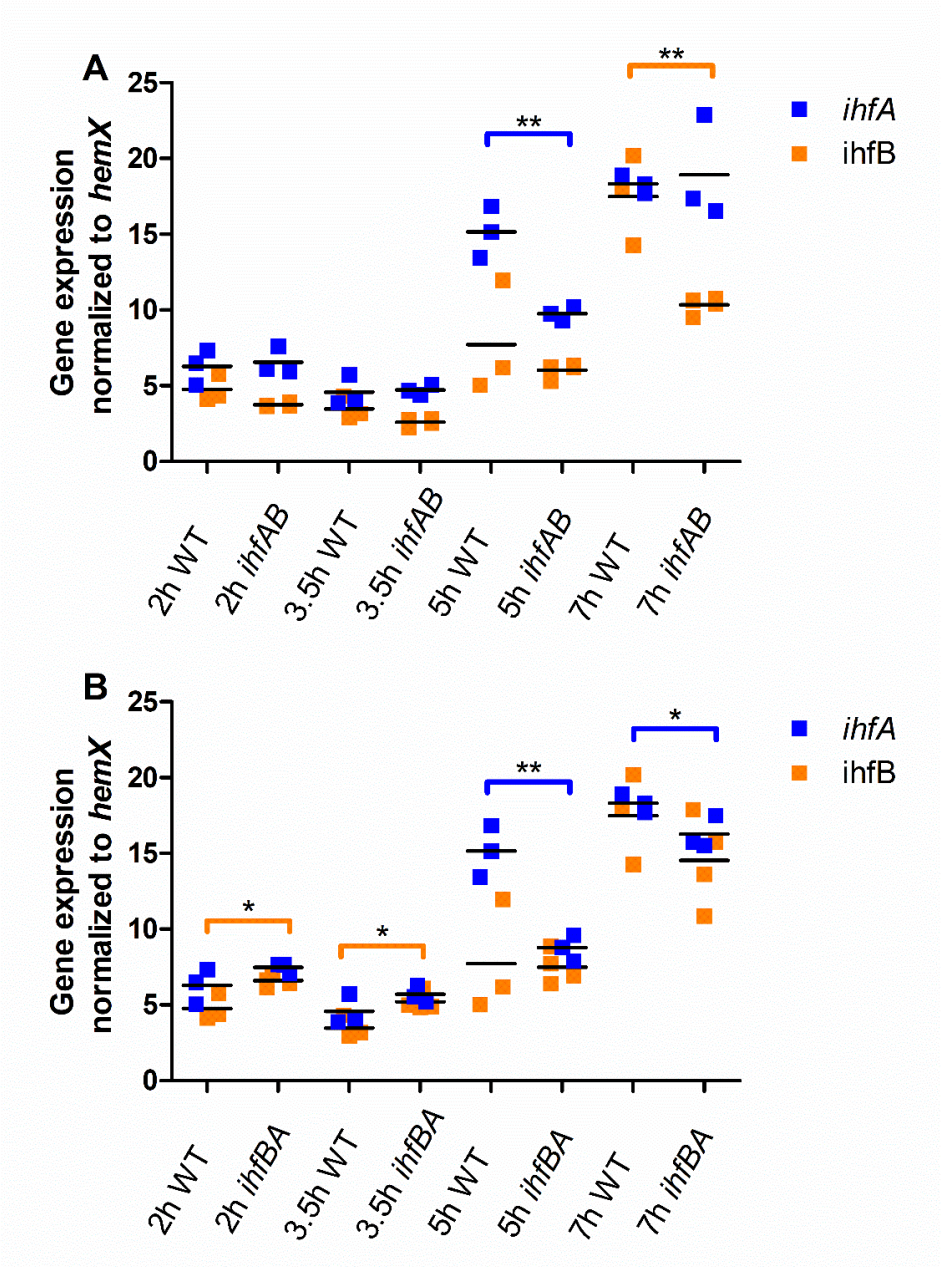


Figure 4.3 Gene expression of *ihfA* and *ihfB* in the SL1344 WT, the *ihfAB* and the *ihfBA* strains. Cells were grown in LB broth at 37°C at 200 rpm and samples were taken at 2 h, 3.5 h, 5 h and 7 h representing lag, exponential, exponential-stationary transition and early stationary growth phases. a) In the *ihfAB* strain, compared to the WT, *ihfA* gene expression decreased at 5 h and peaked at 7 h resembling IHF expression, while *ihfB* gene expression decreased at 5 h and at 7 h. b) In the *ihfBA* strain, compared to the WT, *ihfA and ihfB* gene expression peaked at 7 h resembling IHF expression, *ihfB* gene expression decreased at 5 h and at 7 h. All plots are results of at least three biological replicates, significance was found by Student's unpaired T-test, where $P<0.05$.

4.2.2 Expression of the genes downstream of *ihfA* and *ihfB*

The creation of a strain with chromosomal repositionings, such as the OrfSwap$^{ihfA\text{-}ihfB}$, was associated with a number of uncertainties. For example, it was not clear if the transcription of the created artificial transcription units will be efficiently terminated to avoid an adverse effect on the downstream genes. There is no readthrough transcription at *ihfA* or *ihfB*, according to RNA-seq data (Kröger *et al.*, 2013), indicating that the corresponding terminators are efficient. Nevertheless, to rule out any significant adverse effect, gene expression of the genes downstream from *ihfA* and *ihfB* was monitored in the WT and the OrfSwap$^{ihfA\text{-}ihfB}$ by RT-qPCR.

In *Salmonella* the *btuC* gene, that codes for the vitamin B12 transport system permease, is located 3' to *ihfA* in the WT and 3' to the 5'-*ihfA*[UTR]-*ihfB*[ORF]-3' hybrid locus in the OrfSwap$^{ihfA\text{-}ihfB}$ strain. The *ycaI* gene, that codes for a putative competence-related protein, is located 3' to *ihfB* in the WT and 3' to the 5'-*ihfB*[UTR]-*ihfA*[ORF]-3' hybrid locus in the OrfSwap$^{ihfA\text{-}ihfB}$ strain. Gene expression of both the *btuC* and the *ycaI* was measured by RT-qPCR, as in 2.13, using pairs of the specific primers (SL_btuC_qPCR_Pf and SL_btuC_qPCR_Pr; SL_ycaI_qPCR_Pf and SL_ycaI_qPCR_Pr, Table 2.2) for *btuC* and *ycaI*, respectively. The gene expression of both downstream genes was found to be increased in the OrfSwap$^{ihfA\text{-}ihfB}$, as compared to the WT (Fig. 4.4). Although the increases in gene expression were statistically significant at the most tested timepoints, the resulting gene expression rate was still very low. As seen from the y-axis values (Fig. 4.4), relative gene expression levels of both genes were lower than one, indicating that the gene expression rates remained lower than those of *hemX* – a gene that was used to normalise qPCR data and whose transcription does not change under tested conditions (Kröger *et al.*, 2013). Only a drastic increase in gene expression would have been a cause for concern. This result implies that the exchange of the ORFs did not result in a substantial change in the expression of the downstream genes due to transcriptional readthrough.

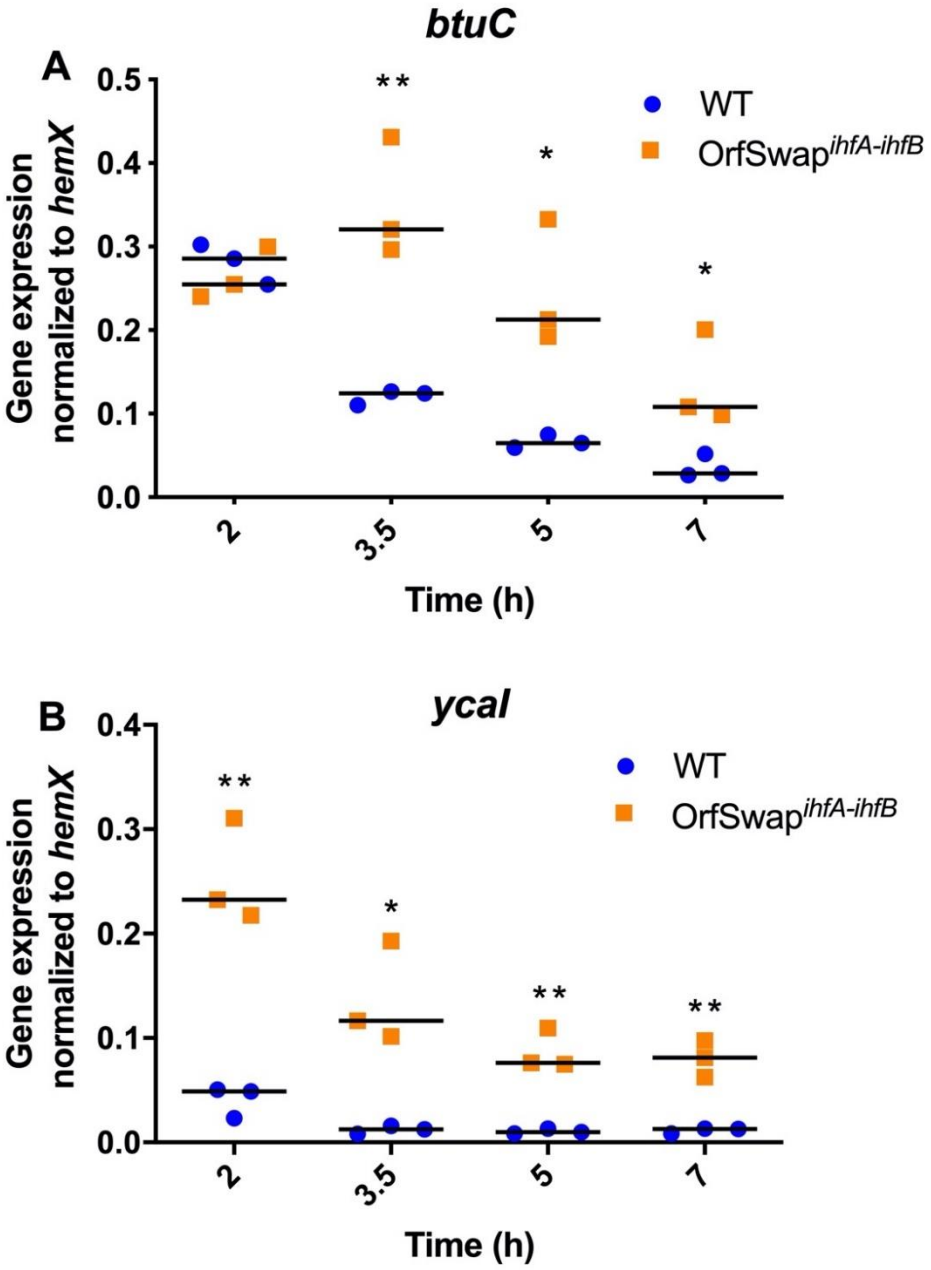


Figure 4.4 Expression of the downstream genes *btuC* and *ycaI* in the WT and the OrfSwap$^{ihfA\text{-}ihfB}$. a) Gene expression of *btuC* – a gene downstream of *ihfA* was measured using RT-qPCR in the WT SL1344 and in the OrfSwap$^{ihfA\text{-}ihfB}$ strains. b) Gene expression of *ycaI* – a gene downstream of *ihfB* was measured using RT-qPCR in the WT SL1344 and in the OrfSwap$^{ihfA\text{-}ihfB}$ strains. The timepoints represent lag, middle exponential, transition from exponential to stationary and early stationary growth phases, respectively. All plots are results three biological replicates. Significance was found by unpaired Student's T-test, where $P<0.05$.

4.2.3 Epitope tagging of *ihfA* and *ihfB*

To complement the gene expression data, IHF protein production dynamics were monitored for each IHF subunit separately across an identical set of timepoints. To achieve this by immunodetection, each IHF subunit was tagged with an epitope that can be readily detected with an antibody.

Tagging of the IhfA and IhfB subunits was performed at the C-terminal end to reduce the chances of a potential adverse effect of the epitope tag on the production or the function of the tagged protein. DNA fragments encoding 3x-FLAG or the 8x-Myc epitope tags were inserted at the 3' end of each *ihf* gene by the Lambda-Red homologous recombination (Datsenko & Wanner, 2000). The tagging of IhfA with the 3x-FLAG tag is described briefly. The 3x-FLAG DNA sequence with an adjacent kanamycin resistance cassette was amplified from the template plasmid pSUB11 with a pair of primers (SL_ihfA_Pf-flag and SL_ihfA_Prev-flag, Table 2.2). The amplicon contained one overhang homologous to the 3' end of the *ihfA* ORF, excluding the stop codon, and the other overhang homologous to the region just downstream from the *ihfA* ORF. The amplicons were transformed into the WT strain carrying plasmid pKD46 and grown in the presence of arabinose, as described in 2.9, deleting the stop codon of *ihfA* (Uzzau *et al.,* 2001). The antibiotic resistance cassette was removed using FLP-mediated site-specific recombination, as described in 2.11 (Cherepanov & Wackernagel, 1995). When required, the 8x-*myc* tag DNA sequence was amplified from the template plasmid pBOP508 (Cho *et al.*, 2006).

Strains in which one of the IhfA or IhfB subunits was tagged with either 3x-FLAG or 8x-Myc in the WT and the OrfSwap$^{ihfA\text{-}ihfB}$ backgrounds were constructed, total of eight strains (Table 2.1). To monitor IhfA and IhfB subunit production in the same strain, double-tagged strains, where each subunit is tagged with a different epitope tag, were constructed. For example, a P22 lysate of the SL1344_*ihfA*::3x-FLAG strain was transduced into the SL1344_*ihfB*::8x-*myc* strain to yield the SL1344_*ihfA*::3x-FLAG *ihfB*::8x-*myc* strain, where *ihfA* and *ihfB* genes were fused to the 3x-FLAG and to the 8x-myc epitope tag sequences, respectively (Table 2.1). Four such differentially tagged

strains were constructed to include all the possible epitope tag combinations (Table 2.1).

4.2.4 The epitope tags do not affect IHF functionality

Epitope tagging may result in an adverse effect on the ability of a protein to fold, assemble and function (Jarvik & Telmer, 1998). The presence of the DNA sequences encoding the tags may also interfere with the expression of the tagged gene. To determine if the 3x-FLAG and the 8x-*myc* sequences had an effect on the gene expression of *ihfA* and *ihfB* in the four double-tagged strains, gene expression of *ihfA* and *ihfB* in those strains was measured by RT-qPCR, as in 4.2.1, and compared to that of the WT and of the OrfSwap$^{ihfA\text{-}ihfB}$. Note that the gene expression of *ihfB* in this section was measured using a different set of qPCR primers (SL_ihfB_qPCR_Pf and SL_ihfB_qPCR_Pr, Table 2.2), thus in the Fig. 4.5b and Fig 4.5d, *ihfB* had a different gene expression magnitude than in the Fig 4.1, Fig. 4.2 and Fig 4.3.

There were statistically significant changes observed between the tagged and the non-tagged strains in many cases (Fig. 4.5). Thus, it was initially concluded that none of the constructed tagged strains should be used for the protein level determination. As the next step, gene expression of *ihfA* and *ihfB* was measured in the strains where only one IHF subunit was tagged with either the 3x-FLAG or the 8x-Myc epitope tag, but the similar significant gene expression changes were also found in those strains. Then the WT and the OrfSwap$^{ihfA\text{-}ihfB}$ strains were tagged with an alternative tag, the hemagglutinin (HA) peptide. HA tag is a small peptide, that is composed of 9 amino-acid residues and derived from influenza virus hemagglutinin HA1. The DNA fragment encoding this tag was expected to be less likely to affect the expression of the tagged genes due to its short length (Schulze-Gahmen *et al.*, 1988). However, during immunodetection the signal was absent from the resulting strains. It was, therefore, decided to return to the use of the double-tagged strains.

In the Fig. 4.5 two out of the four double-tagged strains showed the smallest changes in gene expression. The SL1344_*ihfA*::3x-FLAG *ihfB*::8x-*myc* strain had a lower *ihfA* gene expression level than the WT at 2 h, and a higher gene expression level of *ihfB* than the WT at 7 h. (Fig 4.5a, and Fig. 4.5b). The OrfSwap$^{ihfA\text{-}ihfB}$ _*ihfA*::8x-*myc*_*ihfB*::3x-FLAG strain had *ihfA* and *ihfB* gene expression levels that were identical to those in the untagged strains at all the

tested timepoints (Fig. 4.5c and Fig. 4.5d). Although there were statistically significant changes in gene expression observed in the SL1344_*ihfA*::3x-FLAG *ihfB*::8x-*myc* strain, these were lower in magnitude and less frequent than in the second double-tagged strain in the WT genetic background (Fig 4.5a, and Fig. 4.5b).

Fusing a small gene, such as *ihfA* (300 b.p.-long ORF) and *ihfB* (285 b.p.-long ORF), to the 3x-FLAG (69 b.p.-long) and 8x-*myc* (267 b.p.-long) sequences leads to the production of an mRNA molecule that is significantly longer than *ihfA* or *ihfB* mRNA in the WT. This is important as the longer mRNA has an annealing position for the same pair of qPCR primers altered in relation to its 3' end, and a different degradation pattern compared to the WT mRNA (Hui *et al.*, 2014). Thus, the qPCR results of *ihfA* and *ihfB* in the tagged and non-tagged strains are not directly comparable. Instead, to get an indication of the effect of epitope tagging on the functionality of IHF, it was decided to assess the expression of the genes that are directly regulated by IHF.

The long list of genes regulated by IHF was refined on the basis of conditional expression and regulation (Mangan *et al.*, 2006). To meet the conditional expression requirements, the candidate gene must be expressed at the same growth phase and in the same growth conditions as were used during gene expression measurements. To meet the regulation criterion, IHF must be just one of a few direct regulators of the candidate gene. Three genes: *fis*, *fnr* and *ompR* were chosen to test the effect of the epitope tags on IHF function as they satisfy both criteria. IHF activates *fis*, along with CRP (Pratt *et al.*, 1997); IHF is the sole activator of *fnr* (Mettert & Kiley, 2007); and IHF, along with CRP, is a repressor of *ompR* (Tsui *et al.,* 1991). The correct conditional expression was confirmed with the aid of the original studies that were mentioned above, and with the Salcom –*Salmonella* gene expression compendium (Kröger *et al.*, 2013).

There were no differences in the gene expression of *fis*, *fnr* and *ompR* between the WT and the SL1344_*ihfA*::3x-FLAG *ihfB*::8x-*myc* strains and between the OrfSwap$^{ihfA\text{-}ihfB}$ and the OrfSwap$^{ihfA\text{-}ihfB}$_*ihfA*::8x-*myc*_*ihfB*::3x-FLAG strains at both 2 h and at 7 h, as measured by RT-qPCR using gene-specific primer pairs

(Fig. 4.6), (qPCR primers, Table 2.2). These results indicate that the tagging of IhfA and IhfB with the 3x-FLAG and the 8x-Myc epitope tags did not affect the functionality of IHF. The SL1344_*ihfA*::3x-FLAG *ihfB*::8x-*myc* and the OrfSwap$^{ihfA\text{-}ihfB}$_*ihfA*::8x-*myc*_*ihfB*::3x-FLAG strains were thus used to trace the dynamics of IhfA and IhfB protein production in the WT and the OrfSwap$^{ihfA\text{-}ihfB}$ strains.

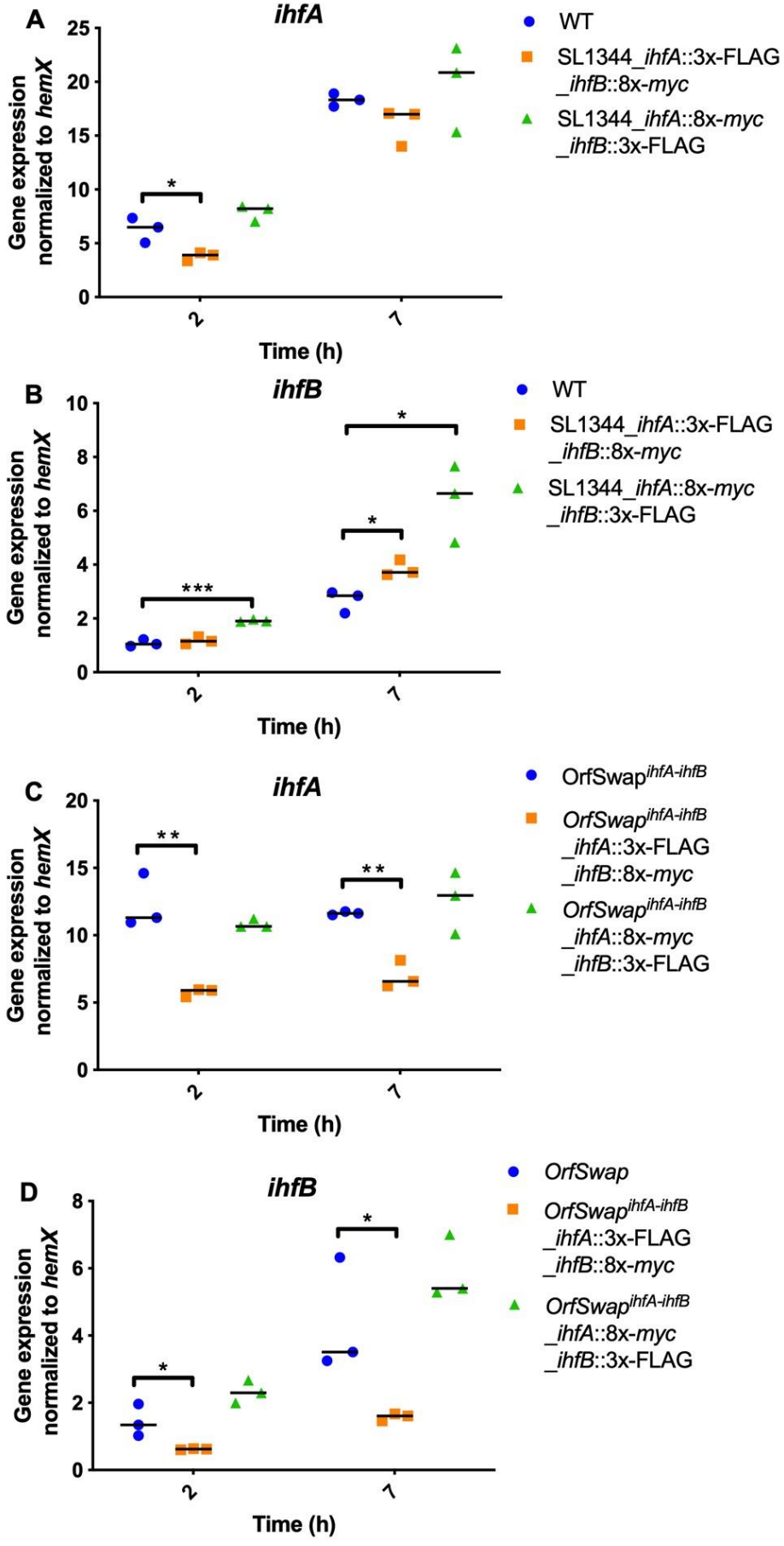

A
ihfA
Gene expression normalized to hemX
Time (h)
WT
SL1344_ihfA::3x-FLAG _ihfB::8x-myc
SL1344_ihfA::8x-myc _ihfB::3x-FLAG
B
ihfB
C
ihfA
OrfSwap ihfA-ihfB
OrfSwap ihfA-ihfB _ihfA::3x-FLAG _ihfB::8x-myc
OrfSwap ihfA-ihfB _ihfA::8x-myc _ihfB::3x-FLAG
D
ihfB
OrfSwap

Figure 4.5 Gene expression of *ihfA* and *ihfB* in the differentially tagged derivatives of the WT and the OrfSwap$^{ihfA\text{-}ihfB}$ strains. Cells were grown in LB broth at 37°C at 200 rpm and samples were taken at 2 h and 7 h representing lag, and early stationary growth phases. a) *IhfA* gene expression was lower at 2 h in the SL1344_*ihfA*::3x-FLAG *ihfB*::8x-*myc* strain than in the WT. b) *IhfB* gene expression was higher at 2 h in the SL1344_*ihfA*::8x-*myc* *ihfB*::3x-FLAG strain and at 7 h in both the SL1344_*ihfA*::3x-FLAG *ihfB*::8x-*myc* and the SL1344_*ihfA*::8x-*myc* *ihfB*::3x-FLAG strains than in the WT. c) *IhfA* gene expression was lower at 2 h and at 7 h in the *OrfSwap*$^{ihfA\text{-}ihfB}$ _*ihfA*::3x-FLAG_*ihfB*::8x-*myc* strain than in the *OrfSwap*$^{ihfA\text{-}ihfB}$. d) *IhfB* gene expression was lower at 2 h and at 7 h in the *OrfSwap*$^{ihfA\text{-}ihfB}$ _*ihfA*::3x-FLAG_*ihfB*::8x-*myc* strain than in the *OrfSwap*$^{ihfA\text{-}ihfB}$. The plots are results of three biological replicates. Significance was found by Student's unpaired T-test, where $P<0.05$.

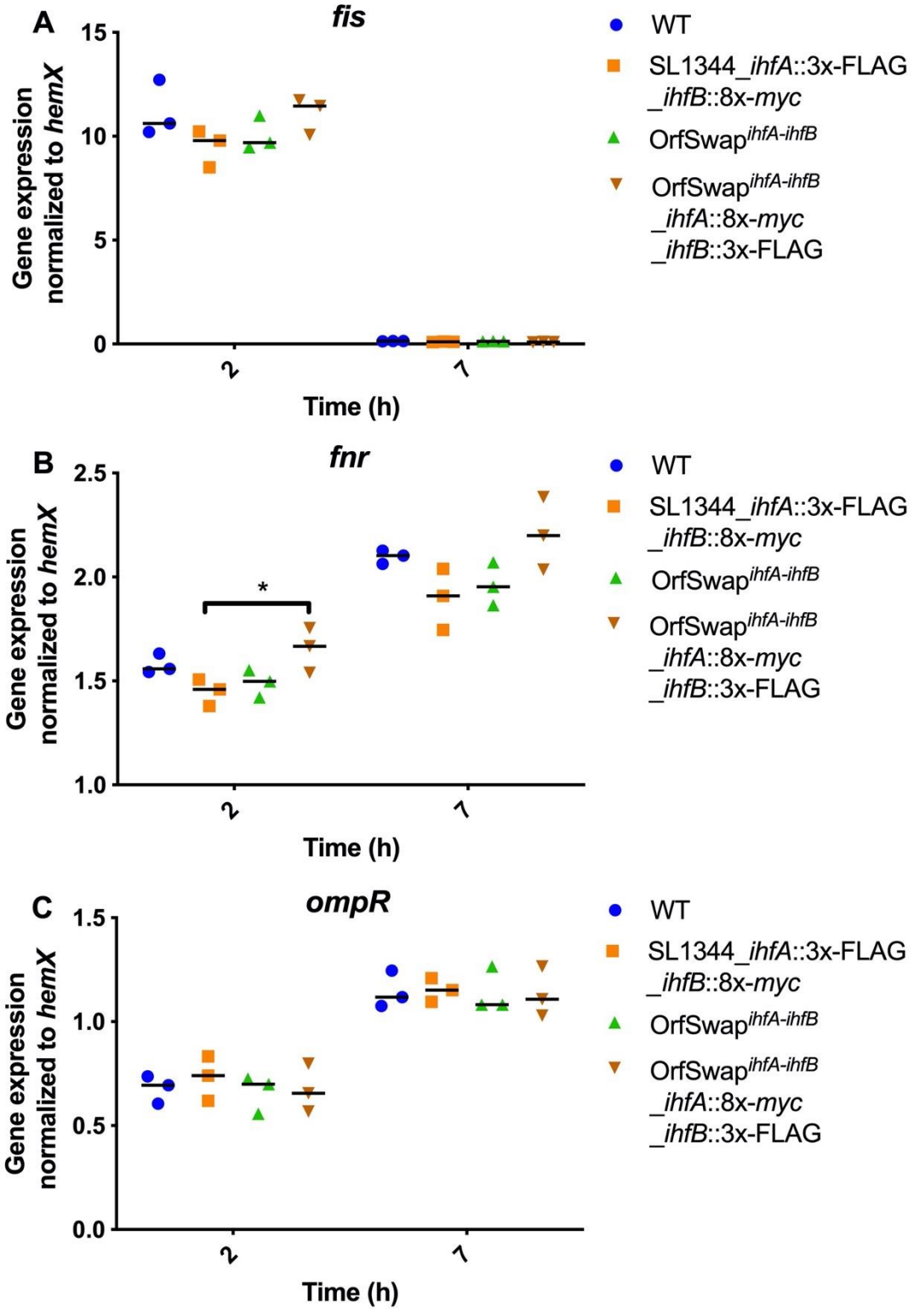

A
fis
Gene expression normalized to hemX
15
10
5
0
2
1
Time (h)
WT
SL1344_ihfA::3x-FLAG _ihfB::8x-myc
OrfSwap ihfA-ihfB
OrfSwap ihfA-ihfB _ihfA::8x-myc _ihfB::3x-FLAG
B
fnr
Gene expression normalized to hemX
2.5
2.0
1.5
1.0
*
2
1
Time (h)
WT
SL1344_ihfA::3x-FLAG _ihfB::8x-myc
OrfSwap ihfA-ihfB
OrfSwap ihfA-ihfB _ihfA::8x-myc _ihfB::3x-FLAG
C
ompR
Gene expression normalized to hemX
1.5
1.0
0.5
0.0
2
1
Time (h)
WT
SL1344_ihfA::3x-FLAG _ihfB::8x-myc
OrfSwap ihfA-ihfB
OrfSwap ihfA-ihfB _ihfA::8x-myc _ihfB::3x-FLAG

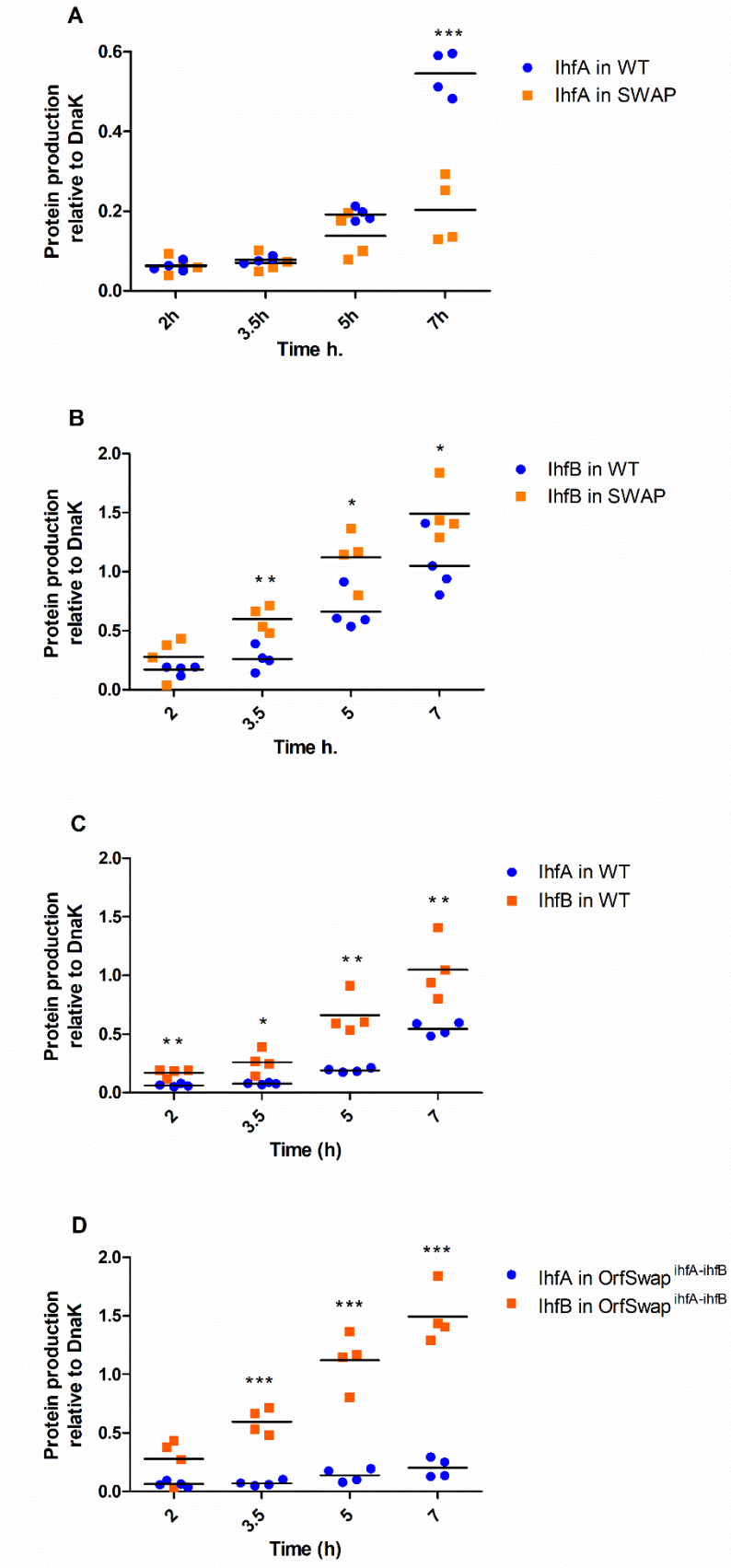
Figure 4.6 Gene expression of *fis*, *fnr* and *ompR* in the differentially tagged strains. Gene expression was compared between the WT, SL1344_*ihfA*::3x-FLAG *ihfB*::8x-*myc*, OrfSwap$^{ihfA\text{-}ihfB}$ and the OrfSwap$^{ihfA\text{-}ihfB}$_*ihfA*::8x-*myc*_*ihfB*::3x-FLAG strains at 2 h and at 7 h, corresponding to the lag at the early stationary growth phases. a) The gene expression levels of *fis* were identical between all the strains within both growth stages. b) The gene expression levels of *fnr* were identical between the strains of the same genetic background at 2 h, and identical between all the strains at 7 h. c) The gene expression levels of *ompR* were identical between all the strains within both timepoints.One-way ANOVA with Tukey's Multiple Comparison Test was used to determine significance, where $p<0.05$.

4.2.5 IhfA and IhfB production is unequal in both the WT and the OrfSwap$^{ihfA\text{-}ihfB}$ strains

The production of the IhfA and IhfB subunits was measured in the double tagged strains SL1344_*ihfA*::3x-FLAG *ihfB*::8x-*myc* and OrfSwap$^{ihfA\text{-}ihfB}$_*ihfA*::8x-*myc*_*ihfB*::3x-FLAG. Bacterial pellets were collected at the same timepoints that were used for the gene expression measurements: 2 h, 3.5 h, 5 h and 7 h. To produce “cleaner” protein lysates where DNA is sheared, cells were lysed by sonication, as in 2.14.1, as opposed to conventional boiled lysates. The samples were run on two SDS-PAGE gels, as in 2.14.4. This is because every sample can be probed simultaneously with anti-FLAG and anti-Myc antibodies. Therefore, after electroblotting the protein samples from the gels onto the separate membranes, as in 2.14.5, one membrane was incubated with anti-FLAG antibody while the other membrane was incubated with anti-Myc antibody, as in 2.14.8, followed by incubation with secondary antibody and visualisation, as in 2.14.9. The typical output is shown in Fig 4.7e and Fig 4.7f, where the corresponding lanes represent the same protein samples.

Densitometry of the imaged membranes was performed with ImageJ software to quantify the results. To compare protein bands from the two membranes, the intensity of each band was divided by the intensity of the corresponding DnaK loading control for normalisation. The results were plotted to illustrate the changes in production of the individual IHF subunits between the WT and the OrfSwap$^{ihfA\text{-}ihfB}$ (Fig. 4.7a and Fig. 4.7b), and to highlight the differences of subunit production within the strains (Fig. 4.7c and 4.7d). Each subunit of IHF displayed an increase in expression at the early stationary growth phase, as expected (Ditto *et al.*, 1994). The production of IhfA was reduced at 7 h (Fig. 4.7a), while the production of IhfB was increased at most timepoints in the OrfSwap$^{ihfA\text{-}ihfB}$ strain (Fig. 4.7b). These changes lead to the increase in the difference between the amounts of IhfA and IhfB in the OrfSwap$^{ihfA\text{-}ihfB}$ strain (Fig. 4.7d). There was considerably more IhfB than IhfA in both the WT and the OrfSwap$^{ihfA\text{-}ihfB}$ strains with the largest difference at 7 h. The lesser amount of IhfA, as compared to IhfB, is not an artefact of the particular epitope tag, since this trend was equally evident with the 3x-FLAG tag (Fig. 4.7e), and with the 8x-Myc tag (Fig. 4.7f). These results show that IhfA and IhfB are not produced in a

1:1 ratio in *Salmonella*, and that the reciprocal exchange of the ORFs of *ihfA* and *ihfB* further increased the difference in the production of the IHF subunits.


A
Protein production relative to DnaK
0.0
0.2
0.4
0.6
2h
3.5h
5h
7h
Time h.
***
IhfA in WT
IhfA in SWAP
B
Protein production relative to DnaK
0.0
0.5
1.0
1.5
2.0
2
3.5
5
7
Time h.
* *
*
*
IhfB in WT
IhfB in SWAP
C
Protein production relative to DnaK
0.0
0.5
1.0
1.5
2.0
2
3.5
5
7
Time (h)
* *
*
* *
* *
IhfA in WT
IhfB in WT
D
Protein production relative to DnaK
0.0
0.5
1.0
1.5
2.0
2
3.5
5
7
Time (h)
***
***
***
IhfA in OrfSwap ihfA-ihfB
IhfB in OrfSwap ihfA-ihfB

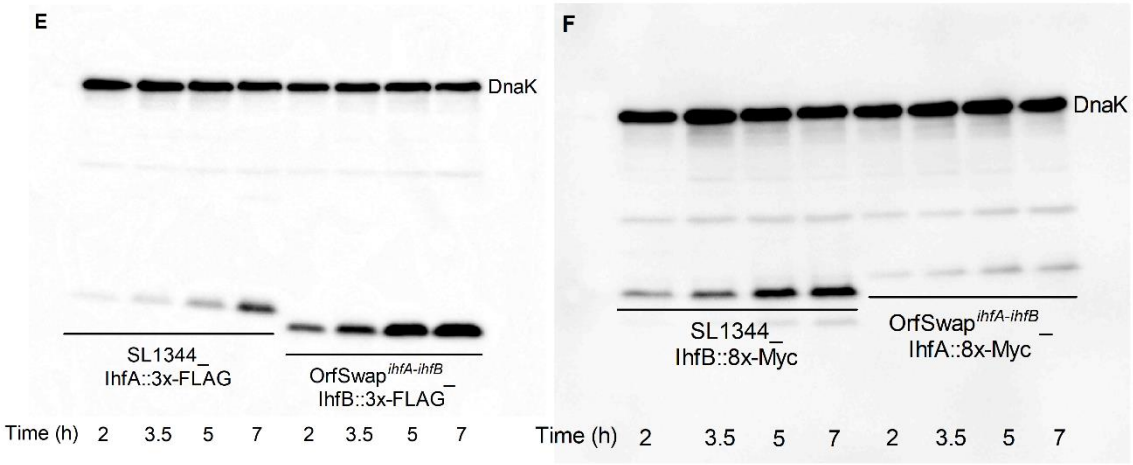


Figure 4.7 IhfA and IhfB subunit production in the WT and the OrfSwap$^{ihfA\text{-}ihfB}$ strains. Densitometry was performed to quantify intensity of bands produced by Western blotting (14.5 kDa IhfA::3x-FLAG, 13.3 kDa IhfB::3x-FLAG, 22 kDa IhfA::8x-Myc, 21 kDa IhfB::8x-Myc, 69.1 kDa DnaK) with normalization against DnaK loading control. Mean of four biological replicates is shown. a) IhfA production in the WT and the OrfSwap$^{ihfA\text{-}ihfB}$. b) IhfB production in the WT and the OrfSwap$^{ihfA\text{-}ihfB}$. c) IhfA and IhfB production in the WT. d) IhfA and IhfB production in the OrfSwap$^{ihfA\text{-}ihfB}$ e) Western immunoblot of the WT (SL1344_*ihfA*::3x-FLAG *ihfB*::8x-*myc*) and the OrfSwap$^{ihfA\text{-}ihfB}$ (OrfSwap$^{ihfA\text{-}ihfB}$_*ihfA*::8x-*myc ihfB*::3x-FLAG) strains probed with 1:20000 anti-FLAG antibody. f) Western immunoblot of the WT (SL1344_*ihfA*::3x-FLAG *ihfB*::8x-*myc*) and the OrfSwap$^{ihfA\text{-}ihfB}$ (OrfSwap$^{ihfA\text{-}ihfB}$_*ihfA*::8x-*myc ihfB*::3x-FLAG) strains probed with 1:20000 anti-Myc antibody. DnaK was used as a loading control, immunoblots are representative of four independent biological replicates, densitometry was done on the four biological replicates, mean is shown. Significance was found by Student's unpaired T-test, where $P<0.05$.

4.2.6 Messenger RNA stability accounts for the relative IHF subunit levels in the WT

While the gene expression measurements by RT-qPCR showed that the *ihfA* and *ihfB* transcripts are present in a cell in close to a 1:1 ratio in both the WT and the OrfSwap$^{ihfA\text{-}ihfB}$ strains (Fig. 4.1), the protein amount estimation by Western blot showed otherwise. According to those data, the amounts of the subunits are almost equal at 3.5 h, however, there is almost twice as much IhfB as IhfA in the WT at 7h – the stage of growth when IHF expression in a cell is at its maximum (Fig. 4.7c). This difference is further exaggerated in the OrfSwap$^{ihfA\text{-}ihfB}$ (Fig. 4.7d). The possibility that differences in *ihfA* and *ihfB* mRNA stabilities might account for the differences in protein levels was investigated.

Rifampicin arrests transcription by binding to the β-subunit of the bacterial RNA polymerase and thus blocking mRNA elongation by steric hindrance (Campbell *et al.*, 2001). Treating a bacterial culture with a transcriptional inhibitor such as rifampicin and then harvesting total RNA at timed intervals is a common method for determining mRNA degradation profiles. RT-qPCR with primers specific for *ihfA* and *ihfB* (Table 2.2) was used as an mRNA detection method (Leclerc *et al.*, 2002). In the WT at 3.5 h, the stabilities of *ihfA* – half-life (T1/2) = 2.78 min and *ihfB* mRNA – T1/2 = 3.54 min were similar (Fig. 4.8a), in agreement with the similar quantities of the subunits seen at this timepoint (Fig. 4.7c). However, at 7 h, *ihfB* – T1/2 = 4.93 min was significantly more stable than *ihfA* mRNA – T1/2 = 2.19 min (Fig. 4.8b), perhaps explaining the almost two-fold excess of IhfB over ihfA at 7 h (Fig 4.7c). The difference between IhfA and IhfB subunits at 3.5 h and especially at 7 h is further increased in the OrfSwap$^{ihfA\text{-}ihfB}$ (Fig.4.7d), and hence, it was expected that measurements of mRNA stabilities will support this. However, in the OrfSwap$^{ihfA\text{-}ihfB}$ strain the relative stabilities of *ihfA* and *ihfB* mRNA were found to be similar, with *ihfA* mRNA becoming slightly more stable than *ihfB.* At 3.5 h T1/2 of *ihfA* mRNA was 2.62 min, while T1/2 of *ihfB* mRNA was 2.32 min (Fig. 4.8c). At 7 h T1/2 of *ihfA* mRNA was 3.38 min, while T1/2 of *ihfB* mRNA was 2.64 min. The half-life of mRNA was determined as described previously (Chen *et al.*, 2008).

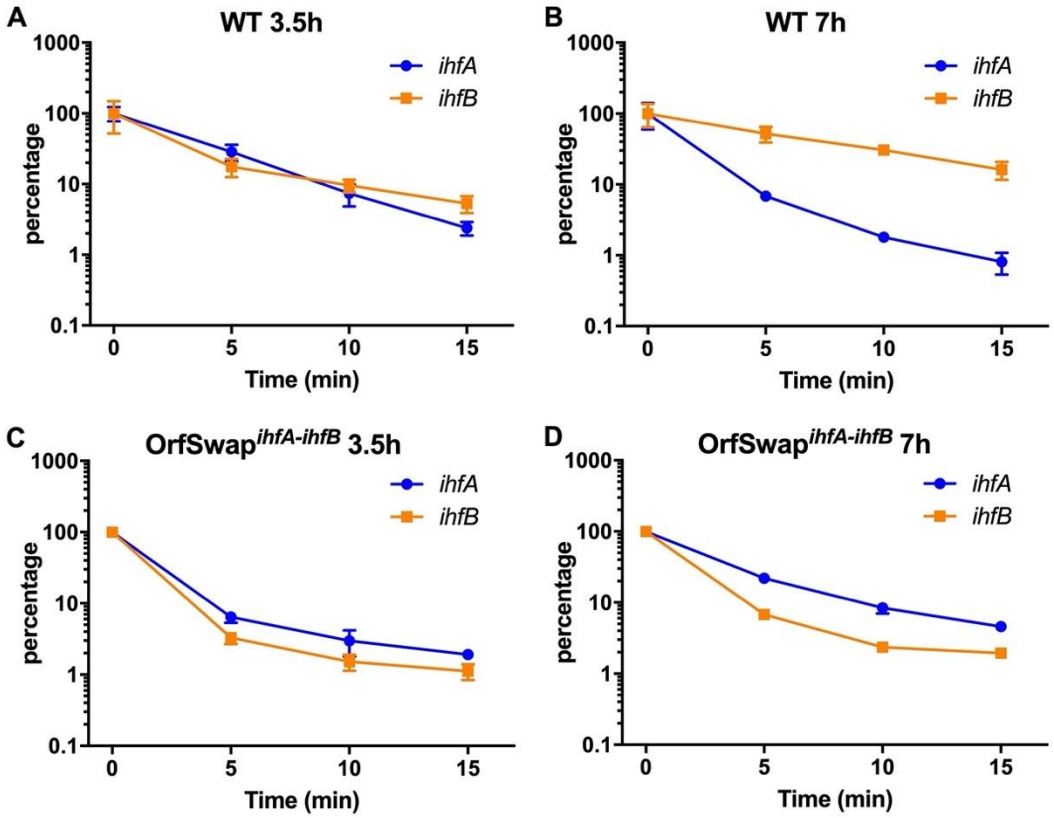


Figure 4.8 *IhfA* and *ihfB* mRNA stabilities in the WT and the OrfSwap$^{ihfA\text{-}ihfB}$ strains. The abundance of mRNA was measured by RT-qPCR before and 5, 10 and 15 min after the addition of 500 µg/ml rifampicin. a) *IhfA* and *ihfB* mRNA stabilities at 3.5 h in the WT. b) *IhfA* and *ihfB* mRNA stabilities at 7 h in the WT. c) *IhfA* and *ihfB* mRNA stabilities a 3.5 h in the OrfSwap$^{ihfA\text{-}ihfB}$ strain. d) *IhfA* and *ihfB* mRNA stabilities at 7 h in the OrfSwap$^{ihfA\text{-}ihfB}$ strain. Mean and standard deviation of three biological replicates is shown.

4.2.7 The global effect of the ORF exchange on the proteome of *Salmonella*

Proteomic study was chosen to assess the impact of the *ihfA* and *ihfB* ORF exchange on the physiology of the cell at a global level. Other types of global studies, such as RNA-seq to determine global RNA production or ChIP-seq to compare global IHF chromosomal binding of the WT and the OrfSwap$^{ihfA\text{-}ihfB}$ would be interesting to perform and compare to the proteomic data, since RNA expression and protein production, although correlate to a certain degree, reveal different aspects of gene expression (Lago *et al.*, 2017). However, it was decided to do proteomics of the WT and the OrfSwap$^{ihfA\text{-}ihfB}$ strains, because it would uncover the alterations in the abundance of proteins – the ultimate products of the gene expression and the molecular machines that do the actual work in a cell. The other reason in favour of proteomic approach is its cost, relative processing ease and availability of the coupled peptide search and statistical analysis software package (Cox *et al.*, 2014; Tyanova *et al.*, 2016). The major disadvantage of proteomics, however, is that it does not provide information directly about small RNAs.

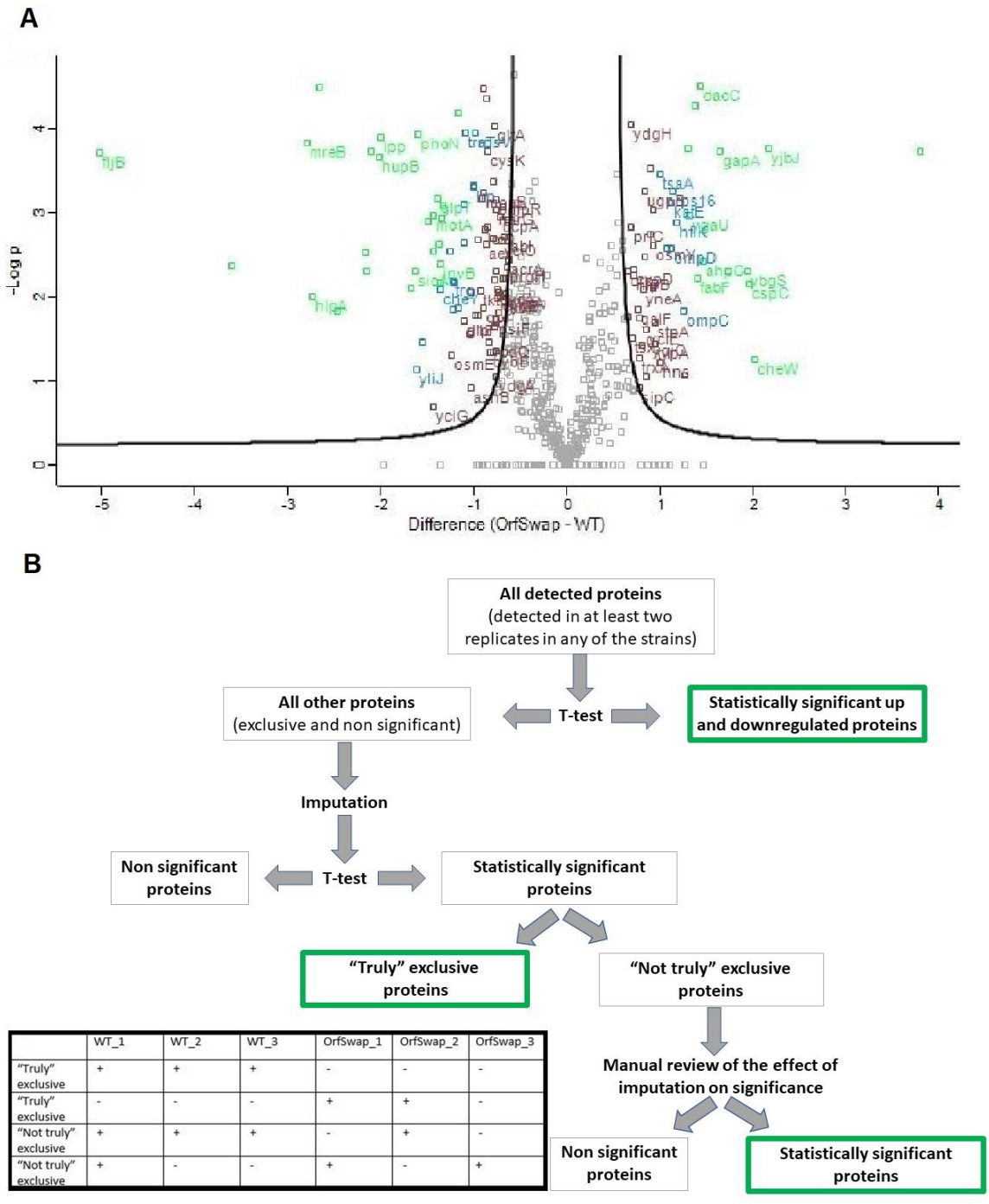
The samples for protein preparation were taken at 7 h of growth in LB to represent the early stationary phase of growth at which IHF is expressed at its highest level and that was one of the time points for the *ihfA* and *ihfB* gene expression measurements by qPCR (section 4.2.1). The protein preparation, Q-Exactive MS runs and the peptide search were performed as described in sections 2.15 and 2.16. Label-free quantification (LFQ) intensity values were used to compare the protein abundance between the samples. These values were obtained by the delayed normalisation algorithm that eliminates differences that arise from separate sample processing and the extraction of maximum peptide ratio information by using only common peptides for the pair-wise sample comparison (Cox *et al.*, 2014). Statistical analysis in Perseus was performed as in section 2.16.3. Two-sided Student's T-test was performed on the data prior to imputation using false discovery rate (FDR) = 0.05 and $S_0 = 1$. This FDR threshold includes proteins with the FDR of 5% and allows, if desired, to increase confidence interval by looking at proteins that have FDR of 1% or 0.1%. The S_0 parameter merges FDR and fold-change and is a standard setting

for Perseus. It is derived from Significance Analysis of Microarrays (SAM) test (Tusher *et al.*, 2001) and results in merging two straight-line thresholds (FDR and fold-change) into smooth hyperbolic curves (Fig. 4.9a). Different S_0 values were tested on the data and S_0=1 was deemed optimal based on the number and the nature of the resulting differentially expressed proteins. The proteins that were differentially expressed at the statistically significant level were filtered out. For all other proteins, missing LFQ intensity values were replaced by imputation from normal distribution. Two-sided Student's T-test was performed on this imputated data using the same FDR and S_0 parameters.

It was decided to distinguish between "truly" exclusive proteins (proteins that were found exclusively in one strain but not in the other) and "not truly" exclusive proteins (proteins that were found in one strain and in one sample only in the other strain) (Fig. 4.9b). Such differentiation was necessary due to the imputation method. It should be noted that while the $\log_2(\text{OrfSwap}^{ihfA\text{-}ihfB})$ – $\log_2(\text{WT})$ difference is real for the up and downregulated proteins, it is an estimate only for the exclusive proteins due to imputation. The "truly" exclusive proteins were not produced in one of the strains at 7 h at the experimental growth conditions. The "not truly" exclusive proteins were detected in a single sample in one of the strains and in two or three samples in the other. Whenever a protein was not detected in one of the strains it was because it was not produced in that strain, due to low expression levels or MS limitations such as inability to detect some peptides or generally poor identification of hydrophobic membrane proteins by MS (Weiner & Li, 2008). Among the "not truly" exclusive proteins, the LFQ intensity value in a single sample was frequently similar to the values of samples in the other strain, however, imputation resulted in the creation of false conclusion about statistical significance. Therefore, all the "not truly exclusive" proteins (the total of 57) were manually reviewed to assess the effect of imputation on statistical significance. The proteins that were clearly affected by imputation were excluded from the list of the unique proteins. As a result, statistically significant up and downregulated, as well as the unique proteins, were deemed differentially expressed.

Principal component analysis (PCA) of the dataset was carried out to ensure that the MS samples cluster according to the strain (Fig. 4.9c), the statistical

analysis was carried out then as above. There were 884 total proteins detected (present in at least two samples in any of the strains). The full output table that includes total, upregulated and downregulated proteins categorised according to Kegg and GO terms is available in appendix I. There were 213 proteins that were significantly downregulated in the OrfSwap$^{ihfA\text{-}ihfB}$ strain. Of these, 34 involved were involved in translation (components of a ribosome or aminoacyl-tRNA biosynthesis), 99 were involved in metabolism (energy, glycolysis, fatty acids, amino-acids metabolism, etc.), 23 proteins were associated with the outer membrane and transport, 16 proteins were involved in genetic information processing, 14 were involved in stress response, 8 proteins were involved in motility and chemotaxis, 7 were involved in pathogenesis, 5 were involved in cytoskeleton and 7 proteins had no known function. In addition, there were 83 proteins that were significantly upregulated in the OrfSwap$^{ihfA\text{-}ihfB}$ strain. Of these, 26 proteins were involved in metabolism, 18 were associated with the outer membrane, 6 were involved in genetic information processing, 5 in translation, 2 in motility and chemotaxis, 2 in cytoskeleton, 2 in pathogenesis, 8 in stress response and 14 proteins had no known function.



	WT_1	WT_2	WT_3	OrfSwap_1	OrfSwap_2	OrfSwap_3
"Truly" exclusive	+	+	+	-	-	-
"Truly" exclusive	-	-	-	+	+	-
"Not truly" exclusive	+	+	+	-	+	-
"Not truly" exclusive	+	-	-	+	-	+

Figure 4.9 Mass spectrometry data of the WT and the OrfSwap$^{ihfA\text{-}ihfB}$. Three biological replicates at 7 h growth in LB. a) Volcano plot of the differentially expressed genes before imputation. The differences in log_2-transformed LFQ intensity were plotted against negative log_{10}-transformed p-values of the two-sided T-test with the $S_0 = 1$. Green represents FDR = 0.001, blue and green represent FDR = 0.01, brown, blue and green represent FDR = 0.05. The analysis was performed in Perseus 1.6.14.0. b) The flow diagram of the strategy used to determine differentially expressed proteins (read from the top). The proteins in green frames were deemed differentially expressed. The insert table defines “truly” exclusive and “not truly” exclusive proteins. “+” represents presence of a protein in a sample, “-“ indicates that a protein was not detected in a sample by MS. c) Principle component analysis of the three WT and OrfSwap$^{ihfA\text{-}ihfB}$ biological replicates shows clustering of the MS samples. The analysis was performed in Perseus 1.6.14.0.

4.2.8 The effect of the ORF exchange on production of proteins associated with processing of genetic information.

The category of proteins referred to as associated with processing of genetic information is broad and involves components of transcription and replication machinery that act on both *Salmonella* chromosome and plasmids, regulatory proteins such as transcription factors and NAPs, topoisomerases, homologous recombination and cell division. Introducing these proteins first is important as they have a direct and an indirect effect on multiple other cellular processes. Such proteins are listed (Fig. 4.10). It should be noted that the $\log_2$ difference values of the “truly” exclusive proteins are estimates only that were obtained by imputation from normal distribution (unfilled bars at Fig. 4.10). These could either be absent from one of the strains or not detected by MS. An example of possible misinterpretation about exclusiveness of a protein is GyrB. It is “truly” exclusive to the WT, but known to be essential for cell survival, thus imputation of exclusive proteins to estimate $\log_2$ difference is a valid way to deal with such proteins. The “truly” exclusive proteins are, therefore, visually distinguished by the unfilled bars at Fig. 4.10.

Mass spectrometry revealed that the exchange of the ORFs that encode IhfA and IhfB resulted in a decrease in the levels of IhfB. This is against the Western blot conclusion that there was an increase of IhfB and a decrease of IhfA in the OrfSwap$^{ihfA\text{-}ihfB}$ (Fig. 4.7a and Fig. 4.7b). It should however be noticed that Western blotting was performed on the epitope-tagged derivatives of the WT and the OrfSwap$^{ihfA\text{-}ihfB}$ strains. While *ihf* gene expression in these strains was unaltered at the transcription level (Fig. 4.5) this was not assessed at the protein production level. The way to do so is to carry out MS analysis of these tagged strains and compare it to the untagged ones (Section 4.2.13). A decrease in β subunit of HU – paralogue of IHF was also seen. The broad regulons of IHF and HU imply that the decreases in levels of IhfB and HupB together may result in a range of alterations in production of multiple other proteins involved in metabolism, virulence, motility, etc. (Mangan *et al.*, 2006; Mangan *et al.*, 2011).

The downregulation of RNA polymerase housekeeping sigma factor RpoD was not expected since RpoD is not a member of the IHF regulon according to previous transcriptomic studies (Mangan *et al.*, 2006; Prieto *et al.*, 2012). However, the effect on RpoD could be indirect since in bacteria IHF is known to frequently carry out its regulatory functions in combination with co-regulators (Colland *et al.*, 2000; Browning *et al.*, 2000; Porter & Dorman, 1997). The downregulation of DNA polymerase I (PolA), single-stranded-DNA binding protein (SSB), plasmid partition protein (ParA) and upregulation of RecA indicates that replication and homologous recombination can be affected in the OrfSwap$^{ihfA\text{-}ihfB}$ strain. However, no differences in growth or morphology between the WT and the OrfSwap$^{ihfA\text{-}ihfB}$ strains were detected (Sections 3.2.3 and 3.2.4). The downregulation of GyrB, TopA, GyrI and MukB suggests that DNA topology and nucleoid superstructure between the strains may vary too. Downregulation of HilA suggests a potential decrease in expression of SPI-1, while downregulation of OmpR may result in an altered response of the OrfSwap$^{ihfA\text{-}ihfB}$ to osmotic and acid stresses (Bajaj *et al.*, 1996; Quinn *et al.*, 2014). Downregulation of the Leucine-responsive regulatory protein (LRP) and upregulation of the cyclic AMP receptor protein (CRP) may result in expression changes of a range of genes that encode many proteins involved in cell membrane composition, transport, virulence and metabolism (Cho *et al.*, 2008; Savery *et al.*, 1996). Upregulation of the translational regulator RNA-binding protein CsrA may have a widespread effect on multiple cellular pathways due to its effect on translation (Romeo *et al.*, 1993). Downregulation of MenG and especially RNAse E – an important endonuclease and a component of degradosome, might decrease the rate of mRNA degradation. Upregulation of H-NS and StpA may result in the altered expression of multiple genes that these proteins regulate via their DNA binding and RNA chaperone activities (Dorman, 2014; Doetsch *et al.*, 2010).

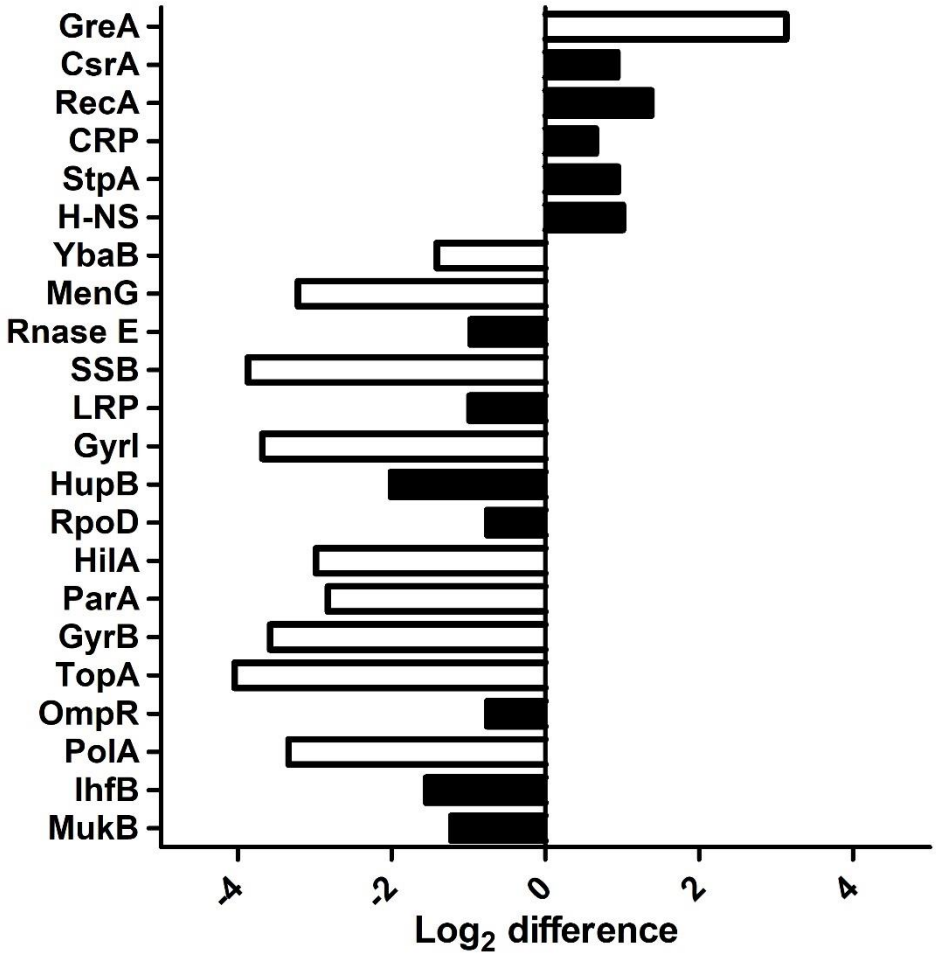


Figure 4.10 Differential expression of proteins associated with processing of genetic information in the OrfSwap*ihfA-ihfB* strain. The log_2 difference between the OrfSwap*ihfA-ihfB* and the WT strains at the early stationary phase of growth. Proteins with positive difference values are upregulated in the OrfSwap*ihfA-ihfB*. Filled bars represent experimental log_2 difference values, unfilled bars represent log_2 difference values that were obtained as a result of imputation.

4.2.9 The effect of the ORF exchange on translation.

Among proteins whose expression was downregulated in the OrfSwap$^{ihfA\text{-}ihfB}$ strain, the largest uniform group was of proteins involved in translation. The main subgroups within were ribosomal proteins and aminoacyl-tRNA synthetases. Most of these proteins were downregulated in the OrfSwap$^{ihfA\text{-}ihfB}$ strain (Fig. 4.11).

The prokaryotic ribosome is made of 30S and 50S subunits that consist of rRNA, small proteins and comprises about one-third protein content of the entire cell (Schuwirth *et al.*, 2005). Downregulation of ribosomal proteins in response to shortage of charged aminoacyl tRNAs is a hallmark of stringent response (Kaczanowska & Rydén-Aulin, 2007; Potrykus & Cashel, 2008) and almost half of all the *Salmonella* ribosomal proteins were downregulated here. Since both, the WT and the OrfSwap$^{ihfA\text{-}ihfB}$ strains were grown in identical conditions in rich medium prior to harvesting for analysis, the possible shortage of amino acids could be a result of amino acid synthesis downregulation in the OrfSwap$^{ihfA\text{-}ihfB}$ strain. The widespread downregulation of translation makes it more difficult to evaluate the results of this proteomic experiment, since less protein synthesis is expected to occur in general. However, although more proteins were downregulated in the OrfSwap$^{ihfA\text{-}ihfB}$ strain, there were 83 upregulated proteins found (Appendix I). This and the necessity to pinpoint the proteomic changes that led to the phenotypes observed in the OrfSwap$^{ihfA\text{-}ihfB}$ strain make it feasible to continue further proteomic investigation but noting the possible impact of the general decrease in proteins related to translation.

Additional complexity in interpreting results of this proteomics experiment arises from the upregulation of CsrA, a global regulator of translation (mentioned above). This protein regulates gene expression post-transcriptionally via its RNA binding properties. It can act both as activator and repressor of translation as well as to attenuate transcription, it can also interact with other global regulatory systems, such as stringent response (Romeo & Babitzke, 2018).

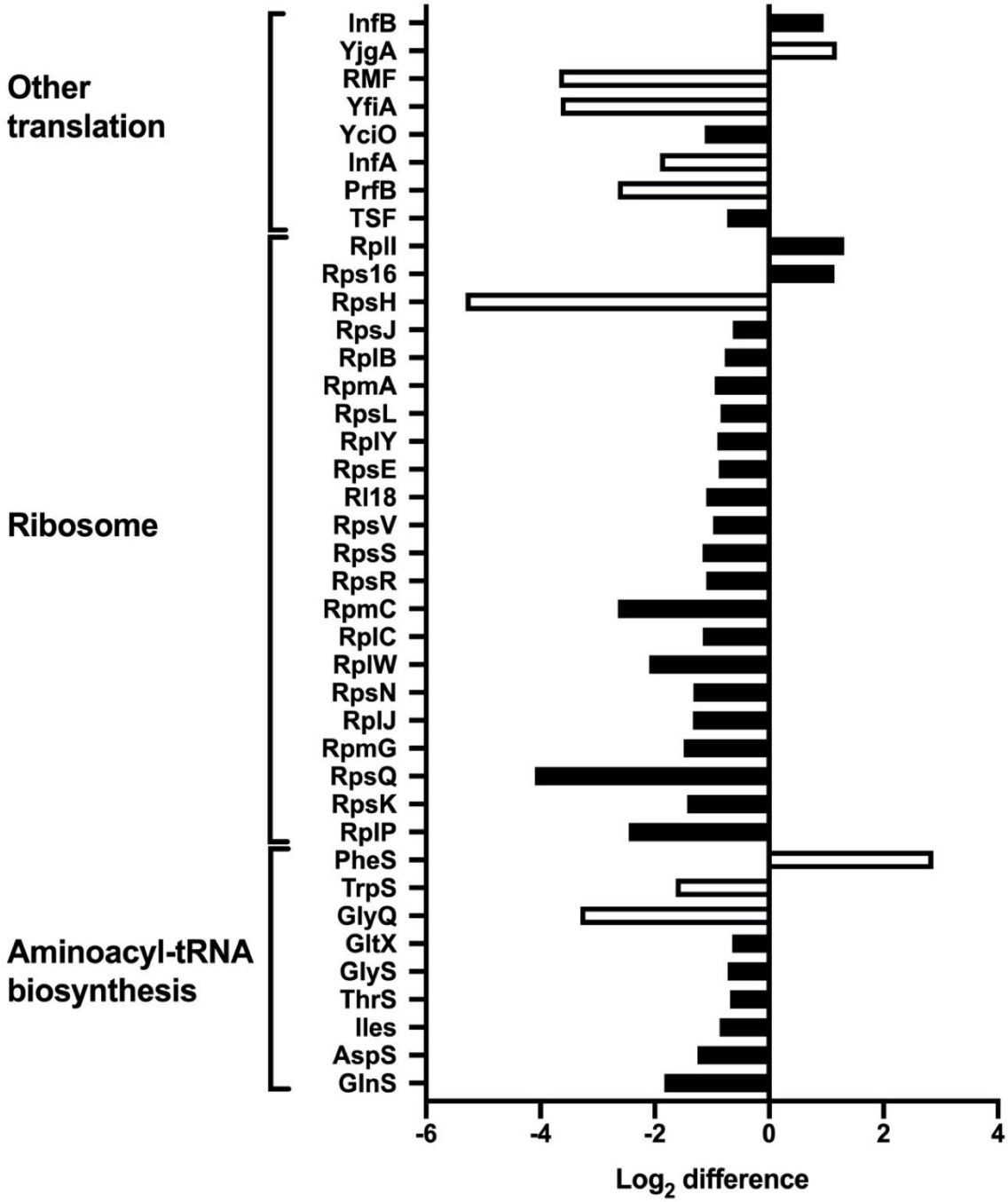


Figure 4.11 Differential expression of proteins associated with translation in the OrfSwap*ihfA-ihfB* strain. The log_2 difference between the OrfSwap*ihfA-ihfB* and the WT strains at the early stationary phase of growth. Proteins with positive difference values are upregulated in the OrfSwap*ihfA-ihfB*. Filled bars represent experimental log_2 difference values, unfilled bars represent log_2 difference values that were obtained as a result of imputation.

4.2.10 The effect of the ORF exchange on metabolism.

IHF is known to regulate multiple metabolic pathways (Mangan *et al.*, 2006). This information was, however, extracted using mutants that lacked one or the other of the IHF subunits, or both. The reciprocal ORF exchange of the IHF-encoding genes is a more delicate genome intervention that had more subtle effect on *Salmonella* than IHF deletions so far.

Nevertheless, multiple metabolic pathways were affected in the OrfSwap$^{ihfA\text{-}ihfB}$ strain (Table 4.1). The negative and positive log_2 difference values represent proteins that are down and upregulated in the OrfSwap$^{ihfA\text{-}ihfB}$ strain, respectively. Note that imputation from normal distribution was used to replace missing LFQ intensity values (section 4.2.7). Therefore, the log_2 difference values of “truly” exclusive proteins are estimates only, and those are marked with (i) in the tables 4.1 and 4.2 The most affected proteins included in the metabolism category were involved in various amino acid synthesis pathways and were downregulated. This is in line with a decrease of the charged aminoacyl-tRNA synthetases and ribosomal proteins noted previously (Section 4.2.9). Most of the other differentially expressed metabolic pathways can be related to the IHF regulon (Mangan *et al.*, 2006; Prieto *et al.*, 2012) or to the regulons of those other global regulators, whose expression was found to be altered in the OrfSwap$^{ihfA\text{-}ihfB}$ strain (Section 4.2.8).

Table 4.1 Differential expression of proteins associated with metabolism in the OrfSwap$^{ihfA\text{-}ihfB}$ strain.

Protein	Log_2 difference	Kegg name
Energy metabolism:		
NarG	-0.7608	Respiratory nitrate reductase 1 alpha chain
CydA	-0.7459	Cytochrome d ubiquinol oxidase subunit I
CyoB	-0.8216	Cytochrome o ubiquinol oxidase subunit I
NuoF	-0.7469	NADH dehydrogenase I chain F
CydB	-0.7062	Cytochrome d ubiquinol oxidase subunit II
MRP	-3.1478 (i)	Conserved hypothetical protein
CysQ	-1.9736 (i)	CysQ protein
NarH	-1.7466 (i)	Respiratory nitrate reductase 1 beta chain
YeiG	-3.5691 (i)	Hypothetical esterase
NuoA	-2.7818 (i)	NADH dehydrogenase I chain A
DmsA	-1.9640 (i)	Anaerobic dimethyl sulfoxide reductase chain A precursor
GltA	-0.7700	Citrate synthase
SucD	0.8991	Succinyl-CoA synthetase alpha chain
NuoI	1.0755	NADH dehydrogenase I chain I
NuoC	0.6670	NADH dehydrogenase I chain C; chain D
Amino acid synthesis:		
AsnB	-1.0909	Asparagine synthetase B
CarB	-3.0612 (i)	Carbamoyl-phosphate synthase large chain
GlmS	-2.8522 (i)	Glucosamine--fructose-6-phosphate aminotransferase
NagA	-1.9806 (i)	N-acetylglucosamine-6-phosphate deacetylase
NanA	-2.7491 (i)	N-acetylneuraminate lyase
NanE	-1.1043	Hypothetical N-acetylmannosamine-6-phosphate 2-epimerase 2 (ec 5.1.3.9) (mannac-6-p epimerase 2)
ProC	-4.1078 (i)	Pyrroline-5-carboxylate reductase
SpeG	-2.9137 (i)	Spermidine N1-acetyltransferase
AstA	-1.2287 (i)	Arginine N-succinyltransferase

Protein	Log_2 difference	Kegg name
GdhA	-2.0004 (i)	NADP-specific glutamate dehydrogenase
CysK	-0.8425	Cysteine synthase A
MetK	-0.6771	S-adenosylmethionine synthetase
MTN	-2.0310 (i)	MTA/SAH nucleosidase
-	-0.9937	Hypothetical glutamate dehydrogenase
GlyA	-2.1570	Serine hydroxymethyltransferase
GcvP	-0.6903	Glycine dehydrogenase [decarboxylating]
SerC	-0.6953	Phosphoserine aminotransferase
ThrC	-1.9614 (i)	Threonine synthase
DapD	-0.7514	2,3,4,5-tetrahydropyridine-2-carboxylate N-succinyltransferase
DapA	-3.6861 (i)	Dihydrodipicolinate synthase
CsiD	-3.1723 (i)	Protein csiD
LuxS	-4.5191 (i)	S-ribosylhomocysteinase
YliJ	-1.9221	Glutathione s-transferase family protein
PepA	-2.1510	Cytosol aminopeptidase
lctD	-0.8689	Hypothetical L-lactate dehydrogenase
DLD	-2.6378 (i)	D-lactate dehydrogenase
BtuE	2.37718 (i)	Hypothetical glutathione peroxidase/vitamin B12 transport periplasmic protein BtuE
-	1.6685 (i)	Hypothetical decarboxylase
DadA	0.8355	D-amino acid dehydrogenase small subunit
KatE	1.1223	Catalase HPII
Fatty acid metabolism:		
FabI	-0.6651	Enoyl-[acyl-carrier-protein] reductase (NADH)
AccA	-3.8962 (i)	Acetyl-coenzyme A carboxylase carboxyl transferase subunit alpha
FabZ	-2.5236 (i)	(3R)-hydroxymyristol acyl carrier protein dehydrase
AccD	-2.5679 (i)	Acetyl-CoA carboxylase beta subunit
FadA	-3.9693 (i)	Small (beta) subunit of the fatty acid-oxidizing multienzyme complex

Protein	Log_2 difference	Kegg name
FabF	1.4100	3-oxoacyl-[acyl-carrier-protein] synthase II
AcpP	3.7919	Acyl carrier protein
FabH	0.6295	3-oxoacyl-[acyl-carrier-protein] synthase III
Glycolisis/ gluconeogenesis:		
PfkA	-1.3680	6-phosphofructokinase
PtsG	-4.1182 (i)	PTS system, glucose-specific IIBC component
AdhC	-3.0160 (i)	Alcohol dehydrogenase class III (ec 1.1.1.1) (glutathione-dependent formaldehyde dehydrogenase) (ec 1.2.1.1) (fdh) (faldh)
YihX	-2.3466 (i)	Hypothetical haloacid dehalogenase-like hydrolase
GlpX	-1.8741 (i)	Hypothetical glycerol metabolic protein
YeaD	2.3937 (i)	Conserved hypothetical protein
GapA	1.6536	Glyceraldehyde 3-phosphate dehydrogenase A
ENO	0.9336	Enolase
LPS biosynthesis:		
KdsA	-0.7188	2-dehydro-3-deoxyphosphooctonate aldolase
LpxA	-2.4774 (i)	Acyl-[acyl-carrier-protein]--UDP-N-acetylglucosamine O-acyltransferase
GmhA	-1.9071 (i)	Phosphoheptose isomerase
Nucleotide metabolism:		
GuaB	-4.1485 (i)	Inosine-5'-monophosphate dehydrogenase
YggV	-2.7739 (i)	HAM1 protein homolog
PurC	-3.1813 (i)	Phosphoribosylaminoimidazole-succinocarboxamide synthase
ThyA	-1.6703 (i)	Thymidylate synthase
PyrB	-1.8611 (i)	Aspartate carbamoyltransferase catalytic subunit
PyrI	-1.7110 (i)	Aspartate carbamoyltransferase regulatory subunit
Pentose phosphate pathway:		
TktB	-0.9261	2 Transketolase

Protein	Log_2 difference	Kegg name
PRS	-0.6285	Ribose-phosphate pyrophosphokinase
KdgK	-3.0764 (i)	2-dehydro-3-deoxygluconokinase
EDA	-1.4803 (i)	KHG/KDPG aldolase
DeoB	0.9044	Phosphopentomutase
Vitamins biosynthesis:		
MoaB	-0.8956	Molybdenum cofactor biosynthesis protein B
SulII	-2.9060	Dihydropteroate synthase type-2 (dihydropteroate synthase type ii) (dhps) (dihydropteroate pyrophosphorylase type ii)
PntB	-1.7726	Pyridine nucleotide transhydrogenase subunit-beta
PhoN	-1.5855	Nonspecific acid phosphatase precursor
IscS	-2.9901	Cysteine desulfurase
PdxB	-3.2608 (i)	Erythronate-4-phosphate dehydrogenase
PdxJ	-2.8779 (i)	Pyridoxal phosphate biosynthetic protein pdxJ
Other carbohydrate metabolism:		
GalK	-1.4289 (i)	Galactokinase
PurU	-2.1582 (i)	Formyltetrahydrofolate deformylase
GPH	-1.5796 (i)	Phosphoglycolate phosphatase
GarRb	-1.3583 (i)	2-hydroxy-3-oxopropionate reductase
GalU	-2.9963 (i)	Glucose-1-phosphate uridylyltransferase
UxaC	-2.0156 (i)	Uronate isomerase
UxuA	-2.4484 (i)	D-mannonate hydrolase
UxuB	-1.3473 (i)	D-mannonate oxidoreductase
GlgC	-2.6771 (i)	Glucose-1-phosphate adenylyltransferase
GlgB	-2.2720 (i)	1,4-alpha-glucan branching enzyme
RfbF	-2.5695 (i)	Glucose-1-phosphate cytidylyltransferase
GlmU	2.0425 (i)	UDP-N-acetylglucosamine pyrophosphorylase
RfbK	1.3272 (i)	Phosphomannomutase
YbhC	1.4866 (i)	Possible pectinesterase precursor
GlgP	1.8931 (i)	Glycogen phosphorylase
GalF	0.7684	UTP-glucose-1-phosphate uridylyltransferase

Protein	Log_2 difference	Kegg name
Other metabolism:		
YadR	-3.1954 (i)	Conserved hypothetical protein
PrpB	-2.7919 (i)	Carboxyvinyl-carboxyphosphonate phosphorylmutase
-	-3.4835 (i)	Conserved hypothetical protein
-	-3.4724 (i)	Conserved hypothetical protein
YciK	-2.4475 (i)	Hypothetical oxidoreductase
YggX	-2.4361 (i)	Conserved hypothetical protein
UcpA	-0.7038	Hypothetical oxidoreductase
YdgJ	-2.5792 (i)	Conserved hypothetical protein
YohF	-2.1044 (i)	Hypothetical oxidoreductase
YdiJ	-2.0867 (i)	Conserved hypothetical FAD-binding protein
SuhB	-1.5299 (i)	Extragenic suppressor protein SuhB
BFR	-3.3178 (i)	Bacterioferritin
GcpE	-3.6692 (i)	4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase
IspB	-3.1139 (i)	Octaprenyl-diphosphate synthase
QOR	-0.7642	Quinone oxidoreductase
DkgA	-0.6499	2,5-diketo-D-gluconic acid reductase A
RlpA	-2.0820 (i)	Rare lipoprotein A precursor
PdgL	-1.6966 (i)	D-alanyl-D-alanine dipeptidase
-	3.2947 (i)	Hypothetical phosphosugar-binding protein
-	2.0759 (i)	Conserved hypothetical protein
TypA	0.953917	GTP-binding protein
DEF	0.864455	Polypeptide deformylase
AhpC	1.4645	Alkyl hydroperoxide reductase c22 protein
TsaA	1.0087	Alkyl hydroperoxide reductase
DacC	1.4333	D-alanyl-D-alanine carboxypeptidase (penicillin-binding protein 6 precursor)

Negative log_2 difference value indicates downregulation in the OrfSwap$^{ihfA\text{-}ihfB}$ strain. (i) indicates estimated values obtained as a result of imputation.

4.2.11 Other proteomic effects of the ORF exchange.

Another large group of proteins that was differentially expressed in the OrfSwap$^{ihfA\text{-}ihfB}$ strain is proteins involved in transport and other components of the outer membrane (Table 4.2). Transmembrane transporters were the only group in the whole proteome that had a larger number of up than downregulated proteins.

Among virulence proteins there were 6 SPI-1 proteins (Table 4.2). This is in line with the observed downregulation of the HilA transcription regulator (section 4.2.8) but overall, only a few virulence proteins were identified since the growth conditions (early stationary phase growth in LB) were not favourable for SPI expression. There were 10 proteins involved in motility and chemotaxis (Table 4.2), including some products of the “middle” genes encoding proteins necessary for flagellar assembly and mostly products of the “late” genes that encoded flagellar components and effector proteins. A motility protein with a $\log_2$ difference of -5.018, the largest obtained without imputation among these (and in the entire proteome) was flagellin FljB. In *Salmonella* flagella can be composed of proteins FljB or FliC, depending on the orientation of the *hin* switch (Bonifield & Hughes, 2003). Substantial downregulation of FljB and the exceeding amount of FliC indicates that in the OrfSwap$^{ihfA\text{-}ihfB}$ strain the *hin* genetic switch is locked in the OFF phase. This effect is not likely to arise directly due to the ORF exchange but is characteristic of *Salmonella* because flagellin is subject to phase variation. Other notable proteins were difficult to categorise because only a few belonged to each category (Table 4.2). An exhaustive list of differentially regulated proteins can be found in Appendix I. Those can be sorted according to the reader’s preferences.

Table 4.2 Differential expression of other notable protein categories in the OrfSwap$^{ihfA\text{-}ihfB}$ strain.

Protein	Log_2 difference	Kegg name
Membrane transporters:		
YhbG	-3.0023 (i)	Probable ABC transport protein, ATP-binding component
ProX	-2.4273 (i)	Glycine betaine-binding periplasmic protein precursor
OppF	-1.3475 (i)	Oligopeptide transport ATP-binding protein (OppF)
GlpT	-1.3775	Glycerol-3-phosphate transporter
TatA	-3.0656 (i)	Sec-independent protein translocase protein
CopA	-2.1426 (i)	Copper-transporting ATPase
GlpF	-1.0977	Glycerol uptake facilitator protein
PutP	-3.4308 (i)	Sodium/proline symporter (proline permease)
CycA	-1.8312 (i)	D-serine/D-alanine/glycine transporter
OmpX	-3.5767 (i)	Outer membrane protein x precursor
YfiO	-2.4147 (i)	Outer membrane protein assembly complex subunit yfiO
HlpA	-2.7315 (i)	Outer membrane protein OmpH precursor
RbsB	4.4376 (i)	D-ribose-binding periplasmic protein
PotD	2.3579 (i)	Spermidine/putrescine-binding periplasmic protein precursor
ArtI	1.8735 (i)	Arginine-binding periplasmic protein 1 precursor
ModA	1.6599 (i)	Molybdate-binding periplasmic protein precursor
MppA	1.8990 (i)	Periplasmic murein peptide-binding protein MppA
PstS	1.3523 (i)	Periplasmic phosphate-binding protein
YobA	2.1862 (i)	Hypothetical cation resistance protein
UgpB	0.8410	Glycerol-3-phosphate-binding periplasmic protein
YneA	0.8262	Hypothetical ABC transport protein, solute-binding component
OsmY	0.9354	Putative periplasmic protein
OmpD	1.1254	Outer membrane protein
NlpB	0.7110	Hypothetical lipoprotein

Protein	Log_2 difference	Kegg name
TSX	0.7143	Nucleoside-specific channel-forming protein tsx precursor
OmpC	1.2658	Outer membrane protein C
Motility and chemotaxis:		
CheY	-1.3523	Chemotaxis protein CheY
CheZ	-2.6895 (i)	Chemotaxis protein CheZ
MotA	-1.4248	Motility protein A
AER	-0.8595	Aerotaxis receptor protein
FljB	-5.0183	Flagellin
FliB	-3.8550 (i)	Lysine-N-methylase (ec 2.1.1.-) (lysine N-methyltransferase)
YmdF	-5.3169 (i)	Conserved hypothetical protein
TRG	-1.2216	Methyl-accepting chemotaxis protein III (mcp-iii) (ribose and galactose chemoreceptor protein)
FliL	1.6068 (i)	Flagella biosynthesis protein
CheW	2.0154	Purine binding chemotaxis protein
Virulence:		
MobC	-1.9280 (i)	Mobilization protein C
PrgI	-3.9180 (i)	Type III secretion system apparatus
PipB2	-3.5652 (i)	Type III secretion system effector protein, Contributes to Sif formation
SicA	-1.6213	Type III secretion-associated chaperone
InvB	-1.3577	Chaperone protein for type III secretion system effectors
PrgH	-0.6347	Type III secretion apparatus component
TlpA	-1.9826 (i)	Alpha-helical coiled coil protein
InvH	2.8418 (i)	Outer membrane lipoprotein
SipA	2.1506 (i)	Pathogenicity island 1 Type III secretion system effector protein
Lipoproteins:		

Protein	Log_2 difference	Kegg name
YbaY	-2.8172 (i)	Conserved hypothetical lipoprotein
YedD	-1.4883 (i)	Hypothetical lipoprotein
LPP	-1.9930	Major outer membrane lipoprotein
YajG	-1.3837	Hypothetical lipoprotein
OsmE	-1.2391	Osmotically inducible lipoprotein E precursor
LolA	2.4213 (i)	Outer membrane lipoprotein carrier protein precursor
YcfM	1.7281 (i)	Hypothetical lipoprotein
Other membrane:		
AcrA	-0.6260	Acriflavin resistance protein A precursor
TraT	-1.0888	Conjugative transfer, surface exclusion protein
-	-0.7692	Hypothetical outer membrane protein
LepB	-1.5898	Signal peptidase I
RcsF	-1.2339	RcsF protein
CopA	-2.1426 (i)	Copper-transporting ATPase
YdgA	3.0338 (i)	Putative transposase
Cytoskeleton:		
MreB	-2.7897	Rod shape-determining protein
DamX	-1.0323	Cytoplasmic membrane protein with SPOR domain
FtsA	-1.4552 (i)	Cell division protein FtsA
StbA	-1.8780 (i)	Plasmid stability protein
ZipA	1.7291 (i)	Cell division protein
FtsZ	0.7902	Cell division protein FtsZ
Chaperons:		
HslU	-0.9538	Heat shock protein
DsbC	-1.3637 (i)	Thiol:disulfide interchange protein
DnaJ	-1.3517	DnaJ protein
FkpA	4.9092 (i)	FKBP-type peptidyl-prolyl isomerase
CypD	0.7222	Peptidyl-prolyl cis-trans isomerase D
SecB	3.6187 (i)	Protein-export protein SecB

Negative log_2 difference value indicates downregulation in the OrfSwap$^{ihfA\text{-}ihfB}$ strain; (i) indicates estimated values obtained as a result of imputation.

4.2.12 Relative quantification of IhfA and IhfB in the WT, OrfSwap$^{ihfA\text{-}ihfB}$ and the tagged derivatives of these strains by Mass Spectrometry.

Along with the relative quantification of proteins across the biological replicates to reveal alterations in the composition of the proteome of the OrfSwap$^{ihfA\text{-}ihfB}$ strain as compared to the WT, relative quantification of the subunits of IHF within each biological replicate was performed. The data generated by the same mass spectrometry experiment were used to determine IhfA to IhfB ratios in the WT and the OrfSwap$^{ihfA\text{-}ihfB}$. The MS was performed separately for the double-tagged derivatives of these strains that were used in Western blotting. MS carried out with the Q-Exactive instrument is a powerful method to accurately determine absolute molar quantities of the proteins (Michalski *et al.*, 2011). This was done to confirm or reject the results related to the IHF stoichiometry obtained with Western blotting in 4.2.5.

Intensity-based absolute quantification values (iBAQ) were used to calculate the ratios of IhfA to IhfB. The iBAQ intensity values are generated by the MaxQuant algorithm that calculates the sum of all peptide peak intensities divided by the number of theoretically observable peptides during tryptic digestion (Schwanhäusser *et al.*, 2011). Absolute molar quantities of all peptides in a sample can be determined if a universal proteomics standard is added to the protein lysate prior to its preparation for the analysis (Arike *et al.*, 2012). Here, however, I was not interested in the molar quantities, but rather in the relative occurrence of IhfA and IhfB, thus no proteomics standard was added.

Followed by the MaxQuant analysis, as in 2.16.2, the protein entries corresponding to IhfA and IhfB were identified in the proteinGroups.txt file. IhfA to IhfB ratios were determined by dividing the corresponding iBAQ intensity values. This was done separately for each biological replicate, the resulting ratios were averaged and plotted (Fig. 4.12). It was found that in the WT the abundance of IhfB was almost two times higher than that of IhfA subunit (0.613 average IhfA/IhfB ratio). This confirms the Western blotting results for the WT that concluded that at 7 h IhfB is present in a cell at the abundance almost

twice as high than that of IhfA (Fig. 4.7c). Interestingly, in the OrfSwap$^{ihfA\text{-}ihfB}$ strain the abundance of IhfB was lower than that on IhfA subunit (2.877 average IhfA/IhfB ratio) (Fig. 4.12). This does not coincide with the Western blotting results for the OrfSwap$^{ihfA\text{-}ihfB}$ strain that concluded that at 7 h IhfB is present in a cell at the higher abundance than that of IhfA and that the difference in abundance between the subunits is increased in the OrfSwap$^{ihfA\text{-}ihfB}$ strain as compared to the WT (Fig. 4.7d).

To investigate the applicability of the MS for quantification of the IHF subunits and to resolve the contradicting conclusions of the MS and the Western blotting, MS of the double-tagged strains that were used in section 4.2.5 for the Western analysis was performed. This experiment was identical to that where the WT and the OrfSwap$^{ihfA\text{-}ihfB}$ strains were analysed and used iBAQ intensity values to determine the relative IhfA and IhfB abundances. It was found that in both the SL1344_*ihfA*::3x-FLAG *ihfB*::8x-*myc* and the OrfSwap$^{ihfA\text{-}ihfB}$_*ihfA*::8x-*myc_ihfB*::3x-FLAG strains the abundance of IhfB was higher than that of IhfA (0.638 and 0.557 IhfA/IhfB ratios, respectively) (Fig. 4.12). This fully confirms the protein production results obtained with Western blotting, showing that IhfB is present in a cell at a higher abundance than that of IhfA in both double-tagged strains and that the difference in abundance between the subunits is increased in the OrfSwap$^{ihfA\text{-}ihfB}$_*ihfA*::8x-*myc_ihfB*::3x-FLAG strain as compared to the SL1344_*ihfA*::3x-FLAG *ihfB*::8x-*myc* strain.

These results indicate that epitope tagging affected IHF expression in the OrfSwap$^{ihfA\text{-}ihfB}$ but not the WT strain. This may seem confusing at first but can be explained by the fact that different combinations of the epitope tags were used for each strain. To reveal other effects due to tagging at the proteome-wide level, the proteomes of the non-tagged and the corresponding double-tagged strains were compared directly.

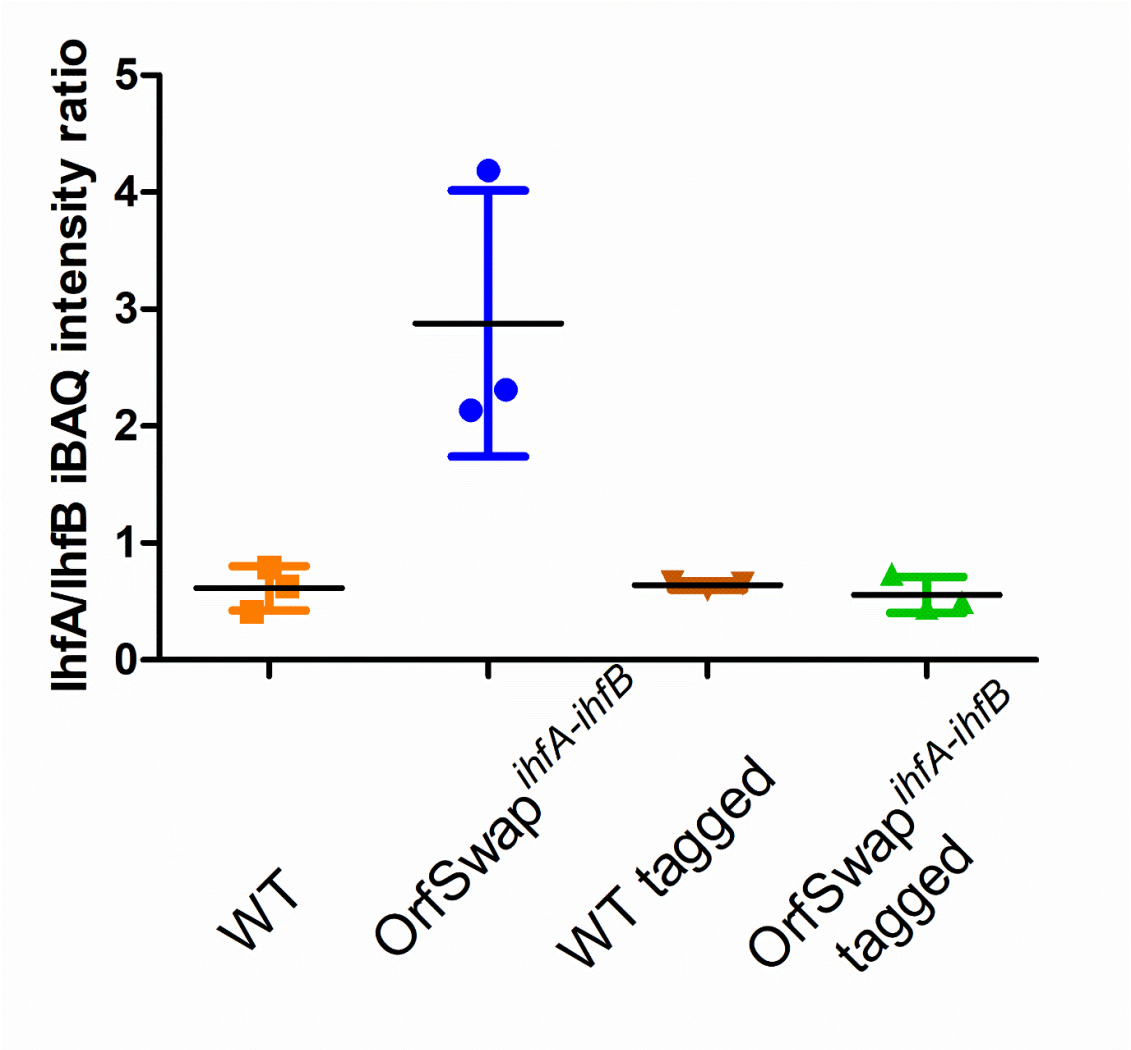


Figure 4.12 IHF subunit ratio in the WT, the OrfSwap$^{ihfA\text{-}ihfB}$, and the tagged strains. Mass spectrometry was performed in the WT, OrfSwap$^{ihfA\text{-}ihfB}$ and their double-tagged derivatives: the SL1344_*ihfA*::3x-FLAG *ihfB*::8x-*myc* and the OrfSwap$^{ihfA\text{-}ihfB}$_*ihfA*::8x-*myc*_*ihfB*::3x-FLAG strains at 7 h, corresponding to the early stationary phase of growth. iBAQ intensity values of the MaxQuant analysis output were used to determine the ratios of IhfA to IhfB subunits in each strain.

4.2.13 Epitope tagging affected the OrfSwap$^{ihfA\text{-}ihfB}$ strain at the proteome level.

The comparison of the WT and the SL1344_*ihfA*::3x-FLAG *ihfB*::8x-*myc* proteomes revealed only three differentially expressed proteins: the methyl-accepting chemotaxis citrate transducer, cell invasion protein SipC and the carbamoyl-phosphate synthase large chain protein CarB were downregulated in the SL1344_*ihfA*::3x-FLAG *ihfB*::8x-*myc*. Such high similarity between the proteomes of these strains suggests that the epitope tagging had minimal effects on the WT strain. This is also supported by the principal component analysis (PCA) (Fig. 4.13a). The lack of the defined clusters corresponding to each strain on the PCA plot is indicative of either the true lack of difference between the strains or problems at the protein preparation stage.

The comparison of the OrfSwap$^{ihfA\text{-}ihfB}$ and the OrfSwap$^{ihfA\text{-}ihfB}$_*ihfA*::8x-*myc*_*ihfB*::3x-FLAG strains revealed multiple differentially expressed and exclusive proteins (Appendix II) and the strain sample groups were clustered separately on the PCA plot (Fig. 4.13b). There were 135 proteins significantly upregulated in the tagged strain. Notably among these, 18 proteins were involved in translation; 57 proteins were involved in metabolism; 14 proteins were involved in the genetic information processing, including sigma factors RpoD and RpoS; 7 proteins were involved in motility and chemotaxis. There also were 79 proteins that were significantly downregulated in the tagged strain. Among these, 10 proteins were involved in genetic information processing, including IhfA and H-NS; 33 proteins were involved in metabolism, 11 proteins were involved in transport and were part of the outer membrane. The full list of differentially expressed proteins can be found in appendix II. The presence of IhfA among the downregulated proteins clearly shows that tagging affected IHF protein production and that this resulted in the misleading conclusions derived from Western blotting that used the OrfSwap$^{ihfA\text{-}ihfB}$_*ihfA*::8x-*myc*_*ihfB*::3x-FLAG strain.

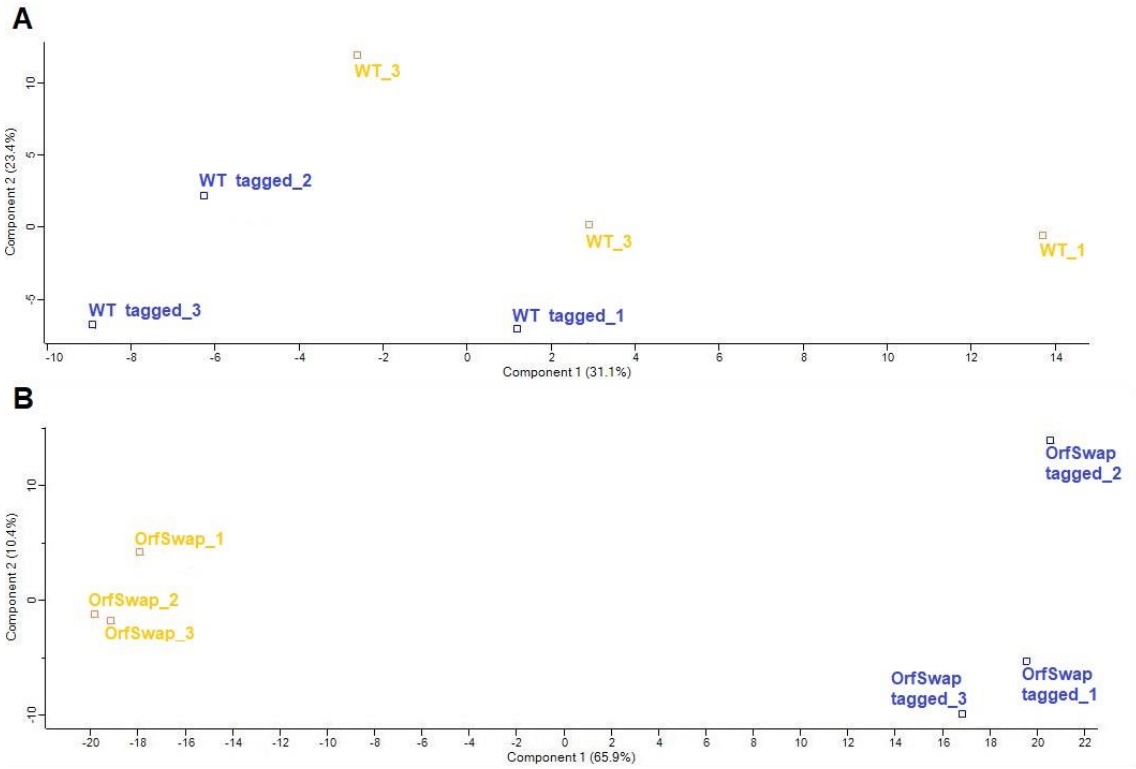


Figure 4.13 Principal Component Analysis (PCA) plots of the untagged and tagged WT and OrfSwap$^{ihfA\text{-}ihfB}$ strains. The identified proteins were clustered based on similarity of LFQ intensity values after imputation. a) Clustering of three biological replicates of the WT and the SL1344_*ihfA*::3x-FLAG *ihfB*::8x-*myc* (WT tagged) mass spectrometry samples. b) Clustering of three biological replicates of the OrfSwap$^{ihfA\text{-}ihfB}$ and the OrfSwap$^{ihfA\text{-}ihfB}$_*ihfA*::8x-*myc*_*ihfB*::3x-FLAG (OrfSwap$^{ihfA\text{-}ihfB}$ tagged) mass spectrometry samples. The analysis was performed in Perseus 1.6.14.0.

4.2.14 The ORF exchange affected the relative concentration of HupA and HupB – subunits of the NAP HU.

HU is a nucleoid associated protein with a structure and characteristics similar to those of IHF. Both proteins belong to the same family of HU-like proteins but have important differences in the extent of DNA path alteration upon binding and in the binding specificity (Swinger & Rice, 2004). HU consists of α and β subunits and can be present in three forms: α_2 homodimer, αβ heterodimer and β_2 homodimer. These conformations are characteristic of the lag, exponential and stationary growth phase, respectively (Claret & Rouviere-Yaniv, 1997).

According to the MS results, section 4.2.8, HupB was found to be downregulated in the OrfSwap$^{ihfA\text{-}ihfB}$ compared to the WT. This result led to further MS investigation that involved determining ratios of the HupA to HupB ratios within the samples with aid of the iBAQ intensity values, as in 4.2.12. This analysis revealed that at 7 h, HupB was more abundant than HupA in the WT strain (HupA/HupB ratio = 0.708) (Fig. 4.14). In the OrfSwap$^{ihfA\text{-}ihfB}$ strain, however, this ratio was inversed (HupA/HupB ratio = 3.594). This shows that the chromosomal exchange of *ihfA* and *ihfB* ORFs led to rescheduling of HupA and HupB protein production. HU protein production dynamics were not additionally assessed due to time constraints.

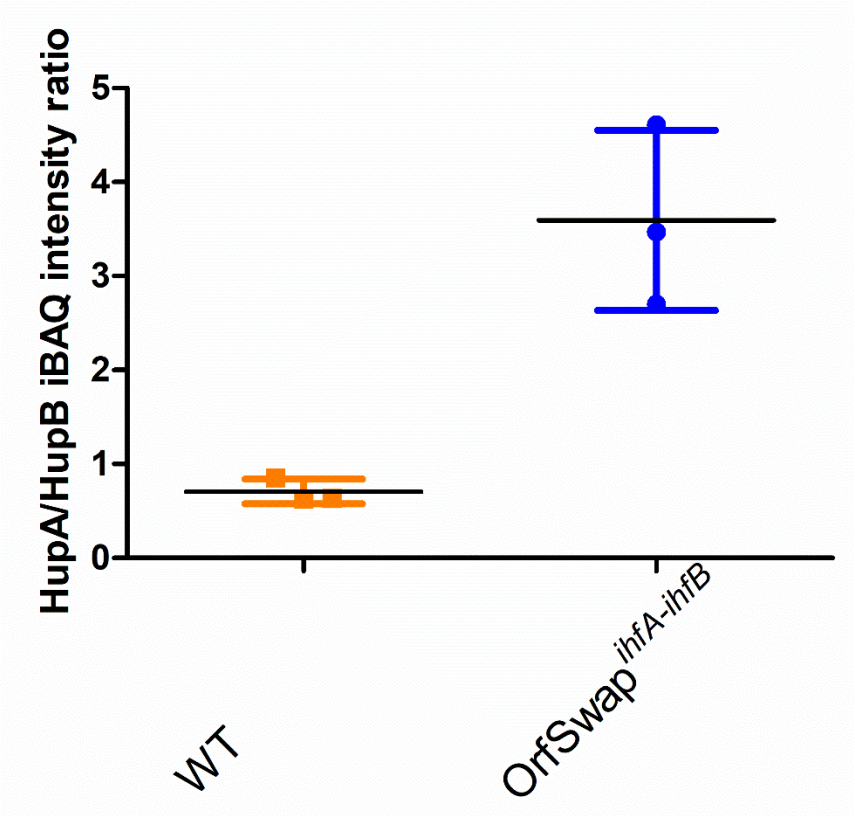


Figure 4.14 HU subunit ratio in the WT and the OrfSwap$^{ihfA\text{-}ihfB}$ strains. Mass spectrometry was performed in the WT and the OrfSwap$^{ihfA\text{-}ihfB}$ strains at 7 h, corresponding to the early stationary phase of growth in three biological replicates. iBAQ intensity values of the MaxQuant analysis output were used to determine the ratios of HupA to HupB subunits in both strains.

4.2.15 Differential expression of RpoD and H-NS is confirmed by Western blotting.

In order to validate the results of mass spectrometry, it was necessary to measure production of proteins that were differentially expressed by alternative method, such as Western blotting. The list of up and downregulated proteins was scanned since MS analysis produced reliable $\log_2$(OrfSwap$^{ihfA\text{-}ihfB}$) – $\log_2$(WT) mean LFQ intensity difference values for such proteins (Section 4.2.8). RpoD and H-NS were chosen because these proteins had the largest mean LFQ intensity difference among proteins against which antibodies were available. In addition, RpoD was downregulated, while H-NS was upregulated in the OrfSwap$^{ihfA\text{-}ihfB}$ strain. The experiment was carried out in two biological replicates as described in Section 2.14.

The Western blotting confirmed the MS results (Fig. 4.15a and 4.15b). According to the MS data, between the two strains there was a $\log_2$ difference of 0.756 for RpoD that corresponds to 1.651-fold difference, while for H-NS there was a $\log_2$ difference of 1.005 that corresponds to 2.127-fold difference. After the target protein bands were quantified by densitometry (Fig. 4.15c) the downregulation of the RpoD and the upregulation of H-NS in the OrfSwap$^{ihfA\text{-}ihfB}$ strain were confirmed. The magnitudes of the changes estimated by Western blotting were much lower than those determined by MS, especially for H-NS. However, this discrepancy in magnitude is not important since Western blotting is not as quantitative as MS and depends a lot on quality and the correct usage of antibodies.

No DnaK loading control was used for these experiments because of the similarity in molecular mass between RpoD and DnaK. Instead, SDS-PAGE was performed using identical amounts of the same protein samples and stained with Coomassie blue (Fig. 4.15d). This demonstrated that equal amounts of total protein had been loaded in each lane. On the SDS-PAGE gel, the WT is characterized by two bands at about 50 kDa (the upper band is indicated by an arrow), while only the lower band of the two is present in the OrfSwap$^{ihfA\text{-}ihfB}$ strain. It is likely that these bands correspond to flagellins that have sizes of 51.611 kDa – FliC and 52.535 kDa – FljB (Liu *et al.*, 2016) and

illustrate downregulation of FljB in the OrfSwap$^{ihfA\text{-}ihfB}$ strain stated in section 4.2.11.

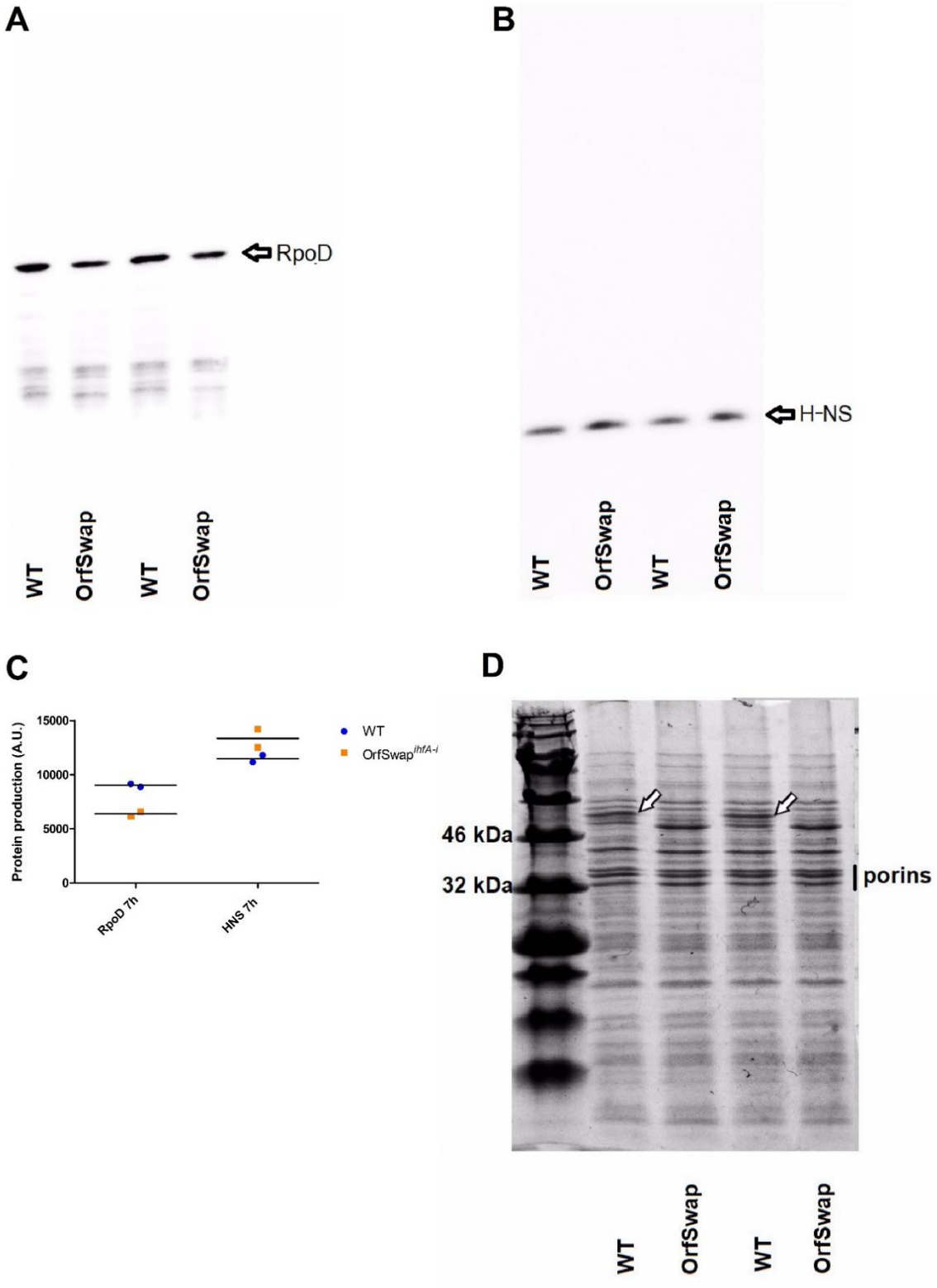

A
RpoD
WT
OrfSwap
WT
OrfSwap
B
H-NS
WT
OrfSwap
WT
OrfSwap
C
Protein production (A.U.)
15000
10000
5000
0
WT
OrfSwap ihfA-i
RpoD 7h
HNS 7h
D
46 kDa
32 kDa
porins
WT
OrfSwap
WT
OrfSwap

Figure 4.15 RpoD and H-NS production in the WT and the OrfSwap$^{ihfA\text{-}ihfB}$ strains. Densitometry was performed in ImageJ to quantify intensity of bands produced by Western blotting (70.5 kDa RpoD, 15.5 kDa H-NS) at 7 h growth in LB. Two biological replicates are shown. a) Western immunoblot of the WT and the OrfSwap$^{ihfA\text{-}ihfB}$ strains probed with 1:20000 anti-RpoD antibody. b) Western immunoblot of the WT and the OrfSwap$^{ihfA\text{-}ihfB}$ strains probed with 1:1000 anti-H-NS antibody c) RpoD and H-NS production in the WT and the OrfSwap$^{ihfA\text{-}ihfB}$ quantified by densitometry of Western blots in a and b. d) SDS page gel of the WT and the OrfSwap$^{ihfA\text{-}ihfB}$ strains stained with Coomassie blue shows equal loading of the samples. Porins OmpA, OmpF and OmpC are highlighted, arrow indicates a band that is present in the WT but absent from the OrfSwap$^{ihfA\text{-}ihfB}$ strain.

4.3 Discussion

Repositionings of the ORFs of *ihfA* and *ihfB* genes resulted in the rescheduled patterns of *ihfA* and *ihfB* gene expression and IhfA and IhfB protein production. The novel gene expression patterns of *ihfA* and *ihfB* in the OrfSwap$^{ihfA\text{-}ihfB}$, *ihfAB*, *ihfBA* and the fact that these patterns are unique for every type of rearrangement is an important result. However, no exchange of the gene expression characteristics was observed when the *ihfA* and *ihfB* ORFs were swapped in this work. This is possibly because *ihfA* and *ihfB* code for the subunits of the same protein that are similarly regulated and do not display individual temporal patterns of expression, as opposed (for example) to the *hupA* and *hupB* or the *fis* and *dps* pairs of genes (Claret & Rouviere-Yaniv, 1997; Ball *et al.*, 1992; Ali Azam *et al.*, 1999). Although the IHF-encoding genes are separately positioned on the chromosome, their locations are proximal to each other and far away from the *oriC* (Fig. 1.1). Thus, the repositionings did not lead to a major change in gene copy number and associated changes of gene expression.

Similarly, in the strains with the artificial *ihf* operons, the ORF that was positioned under the regulatory control of the upstream gene's promoter did not acquire the gene expression characteristics of that upstream gene. Instead, it preserved a gene expression pattern characteristic of itself in the native locus, albeit that the transcript abundance changed (Fig. 4.3). These results suggest that the repositioned IHF-encoding ORFs strive to preserve their original gene expression patterns despite being placed under the influence of the foreign regulatory regions. This could be happening due to the importance of expressing IHF subunits in precise ratios throughout the growth cycle in order to produce fully functional IHF. Preservation of the WT-like gene expression trends in *ihfAB* and *ihfBA* suggests that placing *ihfA* and *ihfB* into operons with each other has milder effects on *Salmonella* than reciprocally exchanging their ORFs.

The protein production dynamics of the IhfA and IhfB subunits were traced separately but within the same strain, to minimise the number of variables that might compromise the validity of the analysis. Strains were created in which the

IhfA and IhfB subunits were tagged with different epitope tags in the same genetic background. The tags were first shown to have no effect on IHF function and then Western blotting of the double-tagged strains was performed. As expected, each subunit of IHF displayed an increase in expression at the early stationary growth phase (Ditto *et al.*, 1994). The IhfA level decreased in the OrfSwap$^{ihfA\text{-}ihfB}$ strain compared to the WT (Fig. 4.7a) and this trend is similar to the one seen in the gene expression of *ihfA* (Fig. 4.2). The IhfB protein production level increased in the OrfSwap$^{ihfA\text{-}ihfB}$ strain compared to the WT (Fig. 4.7b), while the gene expression trend was the opposite (Fig. 4.2). This shows that the repositioning of the genes affected production of the IHF subunits in a way that is not always consistent with the effect on gene expression.

The most surprising result of the protein production study is that the IHF subunits were found not to be expressed in a 1:1 ratio, while it is commonly accepted that the IHF subunits are present in equal amounts in order to assemble into the functional heterodimeric protein (Ditto *et al.*, 1994; Ali Azam *et al.*, 1999). However, the evidence builds that IHF can exist *in vivo* in the form of a homodimer. Each monomeric subunit seems to control its own regulon and this regulon overlaps with that controlled by the other monomer. The individual IhfA and the IhfB regulons also overlap with, but are not identical to, the regulon that is controlled by the IHFAB heterodimer in both *Salmonella* and *E. coli* (Mangan *et al.*, 2006; Prieto *et al.*, 2012). These findings suggest that the IHF subunits have different (but overlapping) transcriptomes and that IHF homodimers can form in the absence of the other subunit. However, while this was shown in the IHF deletion strains, the possibility of a homodimer formation has not been studied in the WT. ChIP-seq was performed in *E. coli* to determine the binding patterns of IhfA and IhfB in a strain with both IHF subunits in place (Prieto *et al.*, 2012). It was shown that the binding signals of IhfA and IhfB were highly similar, but the IhfB covered a significantly higher proportion of the genome than IhfA. Further, in the same study, ChIP-seq was attempted on IhfA and IhfB in the strains with individual IHF subunit deletions but it was not successful, because not enough DNA was recovered for the sequencing part of the experiment. It was therefore hypothesised that

homodimers of IhfB can be formed but only in the presence of a nucleating IHF heterodimer (Prieto *et al.*, 2012). IHF homodimers were shown to form *in vitro*, and the formation of IhfB homodimers was more efficient (Zulianello *et al.*, 1994). My finding that there is almost twice as much IhfB in *Salmonella* than IhfA at the stages of growth when it is most highly expressed (Fig. 4.7) agrees well with the above studies. Further support is provided by the protein abundance database PaxDB (Wang *et al.*, 2015). This is a resource where data from mass spectrometry experiments are compiled and normalized to allow for the direct comparison across species. The amount of IhfB is approximately double the amount of IhfA, according to PaxDB. Thus, data on IhfA and IhfB protein abundance in *S.* Typhimurium were previously available but were not appreciated, as the question of IHF subunit composition was never asked directly before.

The higher levels of IhfB, in spite of the comparable *ihfA* and *ihfB* mRNA synthesis, raises questions about the post-transcriptional fate of *ihfA* and *ihfB* mRNA molecules. The stability of the *ihfA* and *ihfB* transcripts was determined by RT-qPCR of RNA samples extracted from rifampicin-treated cultures. The half-life of *ihfB* mRNA at 7 h was found to be more than twice as long as the half-life of *ihfA* mRNA (Fig. 4.8b). This correlates with almost two times higher level of IhfB subunit at this stage of growth. At 3.5 h, when similar levels of IhfA and IhfB were detected in the WT strain, the half-lives of the corresponding mRNA molecules were similar as well (Fig. 4.8a). Thus, it is proposed that the higher amount of IhfB subunit in the WT *Salmonella* is due to the higher stability of the corresponding mRNA molecule.

Gene expression is extensively controlled at the level of mRNA degradation. The bacterial mRNA degradation depends upon the activity of multiple ribonucleases. Endonucleases, such as RNase E, its homologue RNase G and RNase III, cut RNA molecule internally (Hui *et al.*, 2014). RNase E can act in conjunction with other ribonucleases, as part of a degradosome (Khemici *et al.*, 2008). Exonucleases can be subdivided by their mode of action into the 3' exonucleases, such as PNPase, RNase II and Rnase R , and 5' exonucleases, such as RNase J (Laalami *et al.*, 2014). The discovery of RNase J in *Bacillus subtilis* was the first example of the 5' to 3' proteolytic activity in bacteria (Mathy *et al.*, 2007). To prevent unwanted excessive degradation, mRNA is typically protected by the 5' triphosphate and the 3' stem-loop (Hui *et al.*, 2014), thus

the nature of the 5' and the 3' end of the RNA molecule can determine its degradation rate as many exonucleases degrade mRNA that lacks protection at either end. In addition, the nature of the 5' end can also have an effect on the endolytic cleavage, since RNase E has a preference for the monophosphorylated 5' end, as opposed to the triphosphorylated (Luciano *et al.*, 2017). Based on the above, it is not surprising that an almost reciprocal switch of *ihfA* and *ihfB* mRNA stabilities was observed in the OrfSwap$^{ihfA\text{-}ihfB}$ strain, since in this strain the *ihfA* ORF is under the control of the 5' and 3' untranslated regions of *ihfB* and *vice versa*. Indeed, at 7 h the half-life of the *ihfA* mRNA molecule was found to be longer than that of the *ihfB* mRNA (Fig. 4.8d). The half-lives are, however, inconsistent with the amounts of IhfA and IhfB proteins in the OrfSwap$^{ihfA\text{-}ihfB}$ strain. Perhaps, despite the switch in mRNA stabilities due to the ORF exchange, maintenance of the correct IhfA to IhfB ratio remains of utmost importance for *Salmonella*. Therefore, additional mechanisms, that are possibly related to protein turnover, are operating in the OrfSwap$^{ihfA\text{-}ihfB}$ strain to ensure that the expression of IHF subunits remains optimal for the well-being of the cell. This proposal was proven wrong by further experimentation (see below).

In order to confirm the results of Western blotting with regard to IHF stoichiometry, MS analysis of the WT and the OrfSwap$^{ihfA\text{-}ihfB}$ strains was performed at the stage of growth when the expression of IHF and the difference between the amounts of subunits were maximal, that is at the early stationary phase of growth. It was confirmed that the amount of IhfB is almost twice as high than that of IhfA in the WT (Fig. 4.12). In the OrfSwap$^{ihfA\text{-}ihfB}$ strain, however, an exchange in abundance of the subunits was seen (Fig. 4.12). This contradicted the Western blotting results for the OrfSwap$^{ihfA\text{-}ihfB}$ strain (Fig. 4.7c) but was supported by the mRNA stability investigation (Fig. 4.8d). MS analysis of the strains with the epitope tags that were used in Western blotting was performed next and this confirmed the Western results for those strains (Fig. 4.12). Therefore, the IhfA and IhfB protein abundances were exchanged in response to the exchange of their ORFs in the OrfSwap$^{ihfA\text{-}ihfB}$ strain. The protein production of IhfA and IhfB in the OrfSwap$^{ihfA\text{-}ihfB}$ and the tagged OrfSwap$^{ihfA\text{-}ihfB}$*_ihfA*::8x-*myc_ihfB*::3x-FLAG strains did not match, indicating

that epitope tagging adversely affected the production of IHF. Moreover, a number of changes at the proteome level were detected in the OrfSwap$^{ihfA\text{-}ihfB}$_*ihfA*::8x-*myc*_*ihfB*::3x-FLAG strain (Section 4.2.13). This highlights the fact that a finding no effect of tagging at the level of transcription alone is insufficient to support a claim that epitope tagging does not affect the expression and functionality of the tagged protein (Section 4.2.4). In light of the potential effect of epitope tagging on the ability of a protein to fold, assemble and function (Jarvik & Telmer, 1998), only direct assessment of the production level of the tagged protein and its comparison to that of the WT can confidently indicate the neutrality of an epitope tag.

Apart from the IHF stoichiometric question, MS data were used to assess the global protein production differences that were brought about by the ORF exchange. In previous works the effect of IHF single subunit and double subunit gene deletions on global transcriptomes were studied (Mangan *et al.*, 2006; Prieto *et al.*, 2012), while the genetic manipulations in focus of this work were more subtle and resulted not in the deletion or the disabling of IHF subunits but rather in a complex rewiring of IHF's regulation. This makes placing this work in the context of those earlier studies more difficult. IHF was previously found to regulate multiple genes involved in virulence, motility and chemotaxis, stress and stationary phase-specific genes (Mangan *et al.*, 2006). In the present work, effects on motility and chemotaxis, virulence and stress-related proteins were also seen (Table 4.2) but were not as profound. Some of the same metabolic pathways were also affected (Table 4.1). However, the largest group of affected proteins in the OrfSwap$^{ihfA\text{-}ihfB}$ strain were those associated with translation (Fig. 4.11). In the same previous studies that involved IHF subunit deletion mutants there was a similar number of up- and down-regulated genes, while in this proteomic study there is a clear bias towards downregulation. It is likely to be a result of downregulation of multiple proteins involved in translation that led to fewer proteins being produced in general. It would, thus, be of interest to perform a global transcriptomic study at the same growth conditions to see whether the bias towards downregulation is detected at the level of transcription.

Chapter 5:

Creating a synthetic *gyrBA* operon in *Salmonella enterica* serovar Typhimurium

5.1 Introduction

Genetic material of most bacteria consists of a single circular chromosome, that is not localised to a specific organelle (the nucleus) as in eukaryotic cells. In spite of this seeming simplicity, the bacterial chromosome is a very complex molecule. It is recognised that the genetic material of bacteria does not reside in the cell in a random meaningless conformation but is highly organised in a structure called the nucleoid (Dame *et al.*, 2020). The shape of the nucleoid must at the same time reconcile the needs of two opposing activities – the compaction of the long DNA molecule to fit inside the cell that has a radius of just a few micrometres; and to allow for enzymes that mediate processes such as chromosome replication and segregation, transcription, DNA repair, or gene regulation to access particular sites along the chromosome and genes that need to be expressed or regulated at a particular moment. Thus, the nucleoid shape is highly dynamic, responding to the need of a bacterial cell in constant change of protein production as it grows, engages in cell-to-cell communication, or encounters a change in its environment (Dorman, 2013). The shape of the nucleoid is determined by a variety of factors at multiple levels. These factors include molecular interactions that favour compaction; microdomain organisation determined largely by the activities of nucleoid-associated proteins (NAPs) (Ali Azam *et al.*, 2000; Postow *et al.*, 2004); further organisation into loops of solenoid and plectonemes; and finally macrodomains, assisted by a range of macrodomain-specific proteins (Valens *et al.*, 2004; Dame *et al.*, 2020). The other factors are transcriptional activity (Cabrera *et al.*, 2009; Badrinarayanan *et al.*, 2015) and DNA supercoiling mediated by topoisomerases that can have both local effect at a particular promoter and global effect, making a chromosome negatively supercoiled overall (Dorman & Dorman, 2016).

DNA gyrase is a type II topoisomerase. Its primary function is to introduce negative supercoils into DNA using energy for this process from ATP hydrolysis. The enzyme is a heterotetramer consisting of two GyrA and two GyrB subunits. GyrB is responsible for both interacting with DNA and hydrolysing ATP, while GyrA serves as a scaffold guiding a path for DNA wrapping (Champoux, 2001) (Fig. 1.5). Gyrase subunits highly resemble

subunits of paralogous Topo IV (GyrA aligns with ParC and GyrB with ParE) – another type II ATP-binding topoisomerase that is a decatenase, relaxing positive and negative DNA supercoils (Crisona & Cozzarelli, 2006). Nevertheless, in organisms such as *Mycobacterium* spp, where Topo IV is absent, DNA gyrase is able to provide the decatenase function (Steck & Drlica, 1984; Jain & Nagaraja, 2005). The unique ability of the DNA gyrase among other topoisomerases to introduce negative supercoils makes this enzyme indispensable, since its deletion is not tolerated. The negative DNA supercoiling introduced by DNA gyrase contributes to the overall negatively supercoiled chromosome and is used to compact DNA at many levels of the nucleoid's organisation (Dorman, 2013). DNA gyrase introduces negative supercoils at the *oriC*, enabling replication initiation by DnaA (Samadpour & Merrikh, 2018), while during replication, gyrase removes positive DNA supercoils in front of the moving replication fork (Koster *et al.*, 2010; Kraemer *et al.*, 2019). Positive supercoils are similarly formed in front of the moving RNA polymerase during transcription, creating a substrate for DNA gyrase (Liu & Wang, 1987) (Fig. 1.6). At the local level, it can function as the gene transcription regulator, as supercoiling changes at promoters of many so-called supercoiling-sensitive genes cause them to trigger or stop transcription.

The subunits of DNA gyrase are encoded by *gyrA* and *gyrB* – genes that are differentially regulated and geographically dispersed almost to the opposite poles of *Salmonella's* chromosome (Fig. 1.1). This arrangement does not have an obvious rationale, as co-regulation and expressing both subunits of such an essential enzyme from the same operon at the same locus seems more advantageous. One of the reasons against separate positioning of *gyrA* and *gyrB* is that a gene near the origin will have higher copy number than a gene near the terminus due to the gene dosage effect (Cooper & Helmstetter, 1968). Another caveat is that mRNA is known not to diffuse freely across a cell. It is found to be accumulated at the sites of the protein product requirement (Fei & Sharma, 2018) or tends to be localised near the locus of transcription (Montero Llopis *et al.*, 2010). Synthesis of the subunits of DNA gyrase at sites from where it would be difficult to navigate to each other for assembly into a functional enzyme does not seem a good strategy. Yet, this arrangement is

conserved among gammaproteobacteria (Sobetzko *et al.*, 2012). Interestingly, expression of DNA gyrase subunits from the same locus is not just a theoretical possibility: in many other bacteria, such as *Mycobacterium* spp, *Borrelia* spp and *Streptomyces* spp, DNA gyrase is expressed from a *gyrBA* operon that is located near the origin of chromosomal replication (Unniraman & Nagaraja, 1999; Szafran *et al.*, 2016). The underlying reason for the two different arrangements of *gyrA* and *gyrB* among different bacteria is unclear, but each arrangement is likely to confer a selective advantage in the organism where it is found.

In an attempt to find the significance of the separate positioning of *gyrA* and *gyrB* in *Salmonella*, the *gyrA* ORF was repositioned downstream of *gyrB* to create a *gyrBA* operon, resembling those in *Mycobacterium* and *Streptomyces*. This operon was constructed by a methodology similar to that used when constructing the *ihf* gene operons (Chapter 3), except that the *ihf* operons are completely artificial as they do not occur in nature. The construction and the validation of the *gyrBA* operon, the assessment of the phenotypic and functional outcomes with relevance to DNA gyrase roles and the analysis of the effect of the reposition on transcription are described in detail below. Analysis of the relative *gyrA* and *gyrB* arrangements found in different bacterial species was also carried out.

5.2 Results

5.2.1 Construction of the *gyrBA* strain and an attempt to construct the *gyrAB* strain.

Salmonella with the *gyrBA* artificial operon was constructed using the Lambda-Red homologous recombination (Datsenko & Wanner, 2000) using the same strategy as for the construction of *ihfAB* and *ihfBA* artificial operons, as described in section 3.2.1 (Fig. 3.1). Briefly, a kanamycin resistance cassette was amplified from plasmid pKD4 with a pair of primers (Kan ins gyrA F and Kan ins gyrA R, Table 2.2) using Phusion high-fidelity DNA polymerase. The amplicon had overhangs homologous to a region immediately downstream of *gyrA* and was transformed into the WT strain harbouring plasmid pKD46, grown in presence of arabinose to activate the Lambda Red system, to tag *gyrA* with a kanamycin resistance cassette. The *gyrA*::Kan^R construct, including 20 nucleotides upstream from the *gyrA* starting codon, was amplified using primers (gyrB.int.gyrA::kan_Pf and gyrB.int.gyrA::kan_Prev, Table 2.2) that had both overhangs homologous to sequences immediately downstream of *gyrB.* This was done to enable translation of GyrA protein from the bicistronic *gyrBA* mRNA as several sequences closely matching to a consensus ribosome binding site (5'-AGGAGG-3') are located in this 20 b.p. region. The *gyrA*::kan amplicon was inserted by Lambda Red-mediated recombination immediately downstream of the *gyrB* ORF to construct the *gyrBA* operon. The original *gyrA* gene was deleted by the in-frame insertion of a Kan^R cassette (Baba *et al.*, 2006). All the resistance cassettes were subsequently eliminated via FLP-mediated site-specific recombination (Cherepanov & Wackernagel, 1995). The resulting *gyrBA* strain had the genes that encode both subunits of DNA gyrase being transcribed from the bicistronic operon under the control of a common promoter – P*gyrB* (Table 2.1).

No *gyrAB* operon has been found in other bacteria, so the construction of an artificial *gyrAB* operon in *Salmonella* was proposed and was attempted in parallel with the *gyrBA* construction, using an identical strategy. The *gyrB* gene was tagged with a Kan^R cassette that was amplified using a pair of primers (Kan ins gyrB F and Kan ins gyrB R, Table 2.2). The *gyrB*::Kan^R construct was

amplified using primers (gyrA.int.gyrB::Kan_Pf and gyrA.int.gyrB::Kan_Prev , Table 2.2) and inserted immediately downstream of *gyrA* to construct *gyrAB* operon. To enable translation of the GyrB protein from the bicistronic *gyrAB* mRNA, 20 nucleotides upstream from the *gyrB* translation start codon were included in the insertion, providing the *gyrA* with the best match to a consensus ribosome binding site found there: 5'-ACGAGG-3'. In the resulting strain there was the *gyrAB* operon at the *gyrA* locus and two copies of *gyrB* (Table 2.1). To avoid dealing with outcomes of having increased amount of the GyrB subunit due to the additional copy of *gyrB*, this gene had to be deleted from its original locus. However, repeated attempts to delete *gyrB* from its original locus were not successful. This indicates that the original position of *gyrB* is important, so that it cannot be deleted even though *gyrB* is expressed from the synthetic operon. No further work was performed with the *gyrAB* operon strain.

5.2.2 Strain validation by whole genome sequencing and the need to re-construct *gyrBA*

DNA gyrase is an essential contributor to the global organisation of DNA supercoiling in bacteria. It introduces negative supercoils and removes positive supercoils via a fundamental mechanism that influences the expression of multiple genes and the compaction of DNA. Chromosomal repositioning of genes that encode DNA gyrase may lead to changes in global DNA supercoiling, the expression of supercoiling-sensitive genes, and/or chromosomal compaction. Therefore, compensatory mutations may arise in the strain carrying the translocated gyrase genes to counteract these potentially detrimental changes. Whole genome sequencing was carried out at the Sanger Institute (Hinxton, Cambridgeshire, UK). Sequence contigs were assembled guided by the reference SL1344 sequence NC_016810.1 using de novo assembly with Velvet (Zerbino, 2010). The assembled mutant genome was screened for SNPs by alignment with the same reference using Breseq (Deatherage & Barrick, 2014). A new junction was detected at *gyrB* locus, as expected, indicating successful creation of the *gyrBA* operon. The bioinformatic analysis was kindly performed by Dr. Aalap Mogre.

Apart from the SNPs in *menC* and *manX* that were previously reported (Fitzgerald *et al.*, 2015) and are present in the parental SL1344 WT, there were two other SNPs detected. ACG→GCG SNP, resulting in a T59A substitution, was found in *dxr*, which encodes 1-deoxy-D-xylulose 5-phosphate reductoisomerase. This enzyme catalyses the NADP-dependent reduction of 1-deoxy-D-xylulose-5-phosphate (DXP) to 2-C-methyl-D-erythritol 4-phosphate (MEP) in isopentenyl diphosphate biosynthesis via DXP pathway (Takahashi *et al.*, 1998). Another SNP occurred directly in the artificial *gyrBA* operon and was overlooked during Sanger sequencing – AGA→GGA SNP resulting in an R6G substitution in the repositioned *gyrA* ORF.

Protein-protein interactions of DXR were searched in the STRING database (Szklarczyk *et al.*, 2015), (available at string-db.org) and no connection with DNA gyrase was found. Although DXR is an essential enzyme whose deletion is lethal (Cornish *et al.*, 2006), T59 residue is not a part of enzyme's active

regions, as seen from its crystal structure (Reuter *et al.*, 2002). This suggests that the T59A SNP in *dxr*, is not likely to result in changes at cell physiology level.

The location of the mutation in *gyrBA* is seemingly unimportant – an overhanging end of the N-terminus of *gyrA* that is not involved in any activity and does not even get trapped in crystal structures due to its high mobility (Morais Cabral *et al.*, 1997). However, the amino acid change is significant, as a large positively charged arginine is replaced by a small non-polar glycine. In addition, SNPs in the region that is in the focus of the study should be avoided, so it was decided to construct the *gyrBA* operon strain again.

After the *gyrBA* strain was reconstructed, using identical methodology, whole genome sequencing was carried out at MicrobesNG (Birmingham, UK). There was no SNPs detected, except at *menC* and *manX,* and this strain was used for all other experimentation. In this work the *gyrBA* operon strain with SNPs is referred to as the *gyrBA**, while the reconstructed *gyrBA* operon strain is referred to as the *gyrBA*. In the timeframe of this work, the *gyrBA** was constructed earlier; thus, a number of experiments were carried out first in the *gyrBA**. These experiments had consequently to be repeated in the *gyrBA.* Therefore, for chosen experiments, the *gyrBA* data will be complemented by the data for the *gyrBA**.

5.2.3 Growth characteristics of strains with the *gyrBA* operon.

To evaluate the influence of the *gyrBA* operon creation in *Salmonella* on the growth cycle, the growth patterns of the *gyrBA* strain were measured using absorbance at 600_{nm} and with cell viability counts (Fig. 5.1). No differences were found when growth was measured by both methods, indicating that the growth cycle in *gyrBA* strain is unaltered from that of the WT. Because of this, the 2-h, 3.5-h, 5-h, 7-h and 24-h reference timepoints (that were chosen to represent lag, mid-exponential, exponential-stationary transition, early stationary and late stationary phases of bacterial growth, respectively, in the WT and the strains with repositionings of the IHF-encoding genes) are equally valid in the *gyrBA* strain, and will be used in further experiments.

Before the SNPs in the *gyrBA** strain were detected, its growth profile was also measured by both methods. It was found that the *gyrBA** strain grows significantly slower than the WT between 2 h and 7 h, when measured by OD_{600} (Fig. 5.2a), and between 4 h and 8 h, when measured by viable counts (Fig. 5.2b). The *gyrBA** strain has a delayed exit from the lag phase by about 20 min and lingers behind the WT during the exponential phase, however, later it attains an OD_{600} equal to that of the WT. This small growth defect was thought at first to be caused by the chromosomal repositioning of *gyrA,* prior to the revelation of the *gyrA* and *dxr* SNPs by the genomic sequencing. No growth defect is present in the reconstructed (and SNP-free) *gyrBA* strain.

The growth defect of the *gyrBA** strain must be due to one of its SNPs. To find out which, the *gyrBA** operon was transduced into a genetically clean SL1344 background by P22-mediated generalized transduction to leave behind the SNP in *dxr* while retaining the SNP in the sixth codon of *gyrA*. Growth of the resulting strain was remeasured by absorbance at OD_{600} and was found to be identical to that of the WT (Fig. 5.2c). This shows that T59A SNP in *dxr* was solely causing the growth defect. Given that deletion of *dxr* is lethal (Cornish *et al.*, 2006), mutations in *dxr* may cause growth defects. It is not likely, however, that this mutation was caused directly by the genetic rearrangement since no link of *dxr* to DNA gyrase is known.

Experiments were performed with the WT and the *gyrBA* strains in N-minimal medium with high (10 mM) and low (10 μM) concentrations of Mg^{2+} ions to mimic low Mg^{2+} concentration inside a macrophage. It was necessary to determine timepoints at which the bacterial culture reached mid-exponential and late-stationary growth phases when grown in both types of this minimal medium. To achieve this, growth was measured by absorbance at 600 nm. Pre-conditioned bacterial culture was subcultured into 25 ml of fresh minimal medium of the required Mg^{2+} concentration normalizing to an OD_{600} of 0.03, and OD_{600} values were taken every hour from 2 h until 15 h and at 24 h. To avoid an issue of significant decrease in the culture volume after multiple sampling, two flasks were cultured in parallel and used for measurements between 2 h and 8 h. Separate flasks were similarly set up for measurements between 8 h and 15 h. Growth dynamics of both the WT and the *gyrBA* strains were found to be identical at every Mg^{2+} concentration. At high Mg^{2+} both strains reached higher OD_{600} values. At low Mg^{2+} both strains exited the lag phase more rapidly but entered the stationary phase much earlier and achieved lower OD_{600} values (Fig. 5.3). Mid-exponential phase was estimated to be located at 7 h for low Mg^{2+} cultures and at 9 h for high Mg^{2+} cultures. The 24-h timepoint was chosen to represent late-stationary phase of growth. Minimal medium is more demanding for bacteria than rich medium, therefore, in minimal medium weak growth defects, that are undetectable in rich medium, are more likely to manifest themselves. Identical growth patterns of the WT and the *gyrBA* strains in N-minimal medium highlight that expressing DNA gyrase from the *gyrBA* operon made no effect on the growth cycle (and hence the cell cycle) of *Salmonella*.

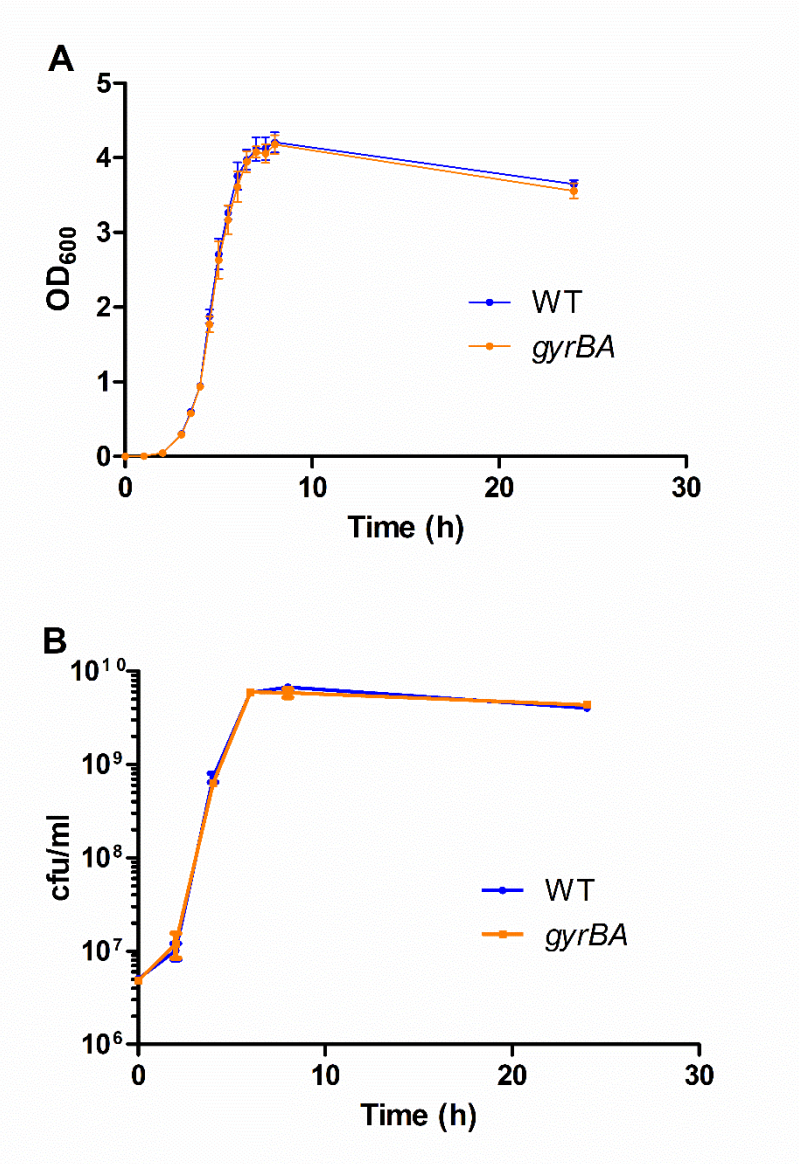


Figure 5.1 Growth characteristics of the *gyrBA* strain. a) Growth of the *gyrBA* strain as measured by absorbance at 600_{nm}. OD_{600} values were taken every hour until 3 h, then every 30 min until 8 h and lastly at 24 h. b) Growth of the *gyrBA* strain as measured by viability counts. Dilutions of bacterial cultures were spread on agar plates, incubated at 37°C and colonies were counted. All plots are results of at least three biological replicates, error bars represent standard deviation.

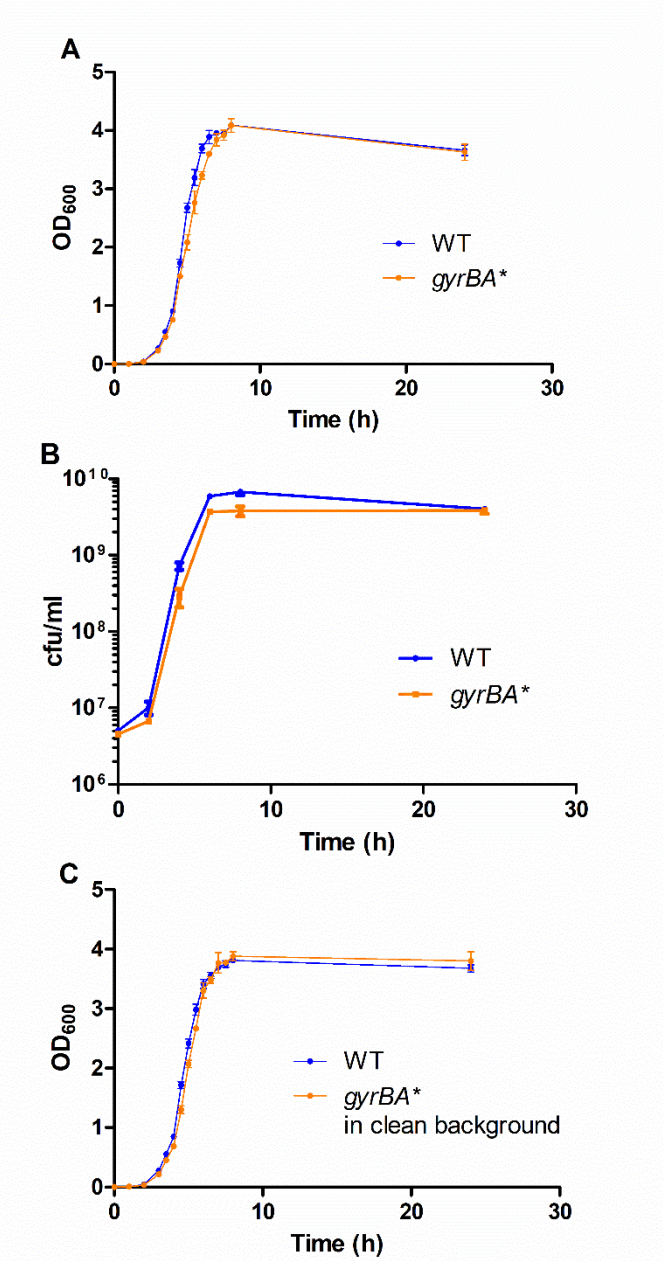


Figure 5.2 Growth characteristics of the *gyrBA strain.** a) Growth of the *gyrBA** as measured by absorbance at 600_{nm}. The difference is significant between 2 h and 7 h. b) Growth of the *gyrBA** as measured by viability counts. The difference is significant between 4 h and 8 h. c) Growth of the *gyrBA** transduced into a clean genetic background as measured by absorbance at 600_{nm}. All plots are results of at least three biological replicates, error bars represent standard deviation. Significance was found by unpaired Student's T-test when $p<0.05$.

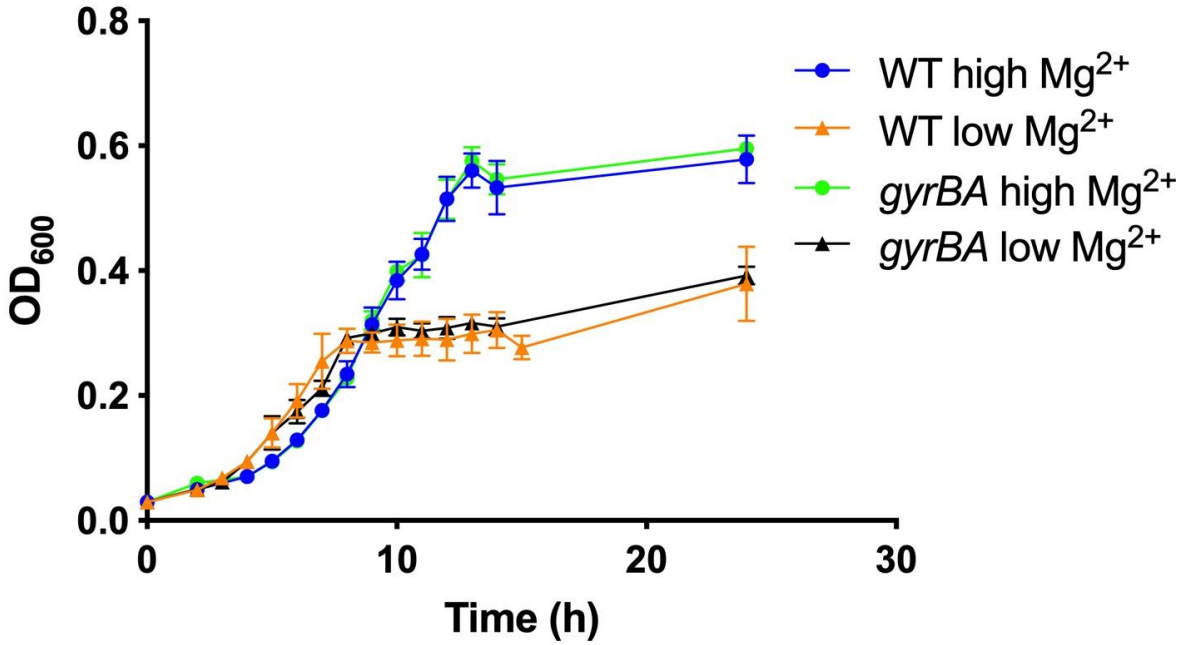


Figure 5.3 Growth characteristics of the WT and the *gyrBA* strains in minimal medium with high and low concentrations of Mg^{2+} ions. Growth of the WT and the *gyrBA* strains as measured by absorbance at 600 nm. Pre-conditioned culture was subcultured into 25 ml of fresh minimal medium of the required Mg^{2+} concentration, normalizing to an OD_{600} of 0.03. OD_{600} values were taken every hour from 2 h until 15 h and at 24 h. All plots are results of at least three biological replicates, error bars represent standard deviation.

5.2.4 Cell morphology of *Salmonella* with the *gyrBA* operon

DNA gyrase has a direct role in the cell cycle. During the initiation of chromosomal replication, DNA gyrase introduces negative supercoils at the *oriC*, enabling replication initiation by DnaA (Samadpour & Merrikh, 2018). DNA gyrase is also able to decatenate sister chromosomes (Steck & Drlica, 1984); although this function is most often provided by another type II topoisomerase – topoisomerase IV (Espeli *et al.*, 2003). Apart from that, DNA gyrase takes part in the cell cycle indirectly, by resolving positive supercoils that form in front of the moving replication fork, enabling chromosomal replication (Koster *et al.*, 2010). Thus, in theory, if the *gyrBA* operon had expressed DNA gyrase that is less processive than the WT, changes in cell morphology would have been observed.

Cultures of the *gyrBA* and the *gyrBA** were grown to the mid-exponential and the late-stationary growth phases and observed by phase-contrast microscopy, comparing to the WT morphology (Fig. 5.4). There were no deviations from the WT *Salmonella* morphology. This indicates that even if DNA gyrase, that is synthesised from the operon, differs from its WT counterpart, these differences are not sufficient to cause morphological changes.

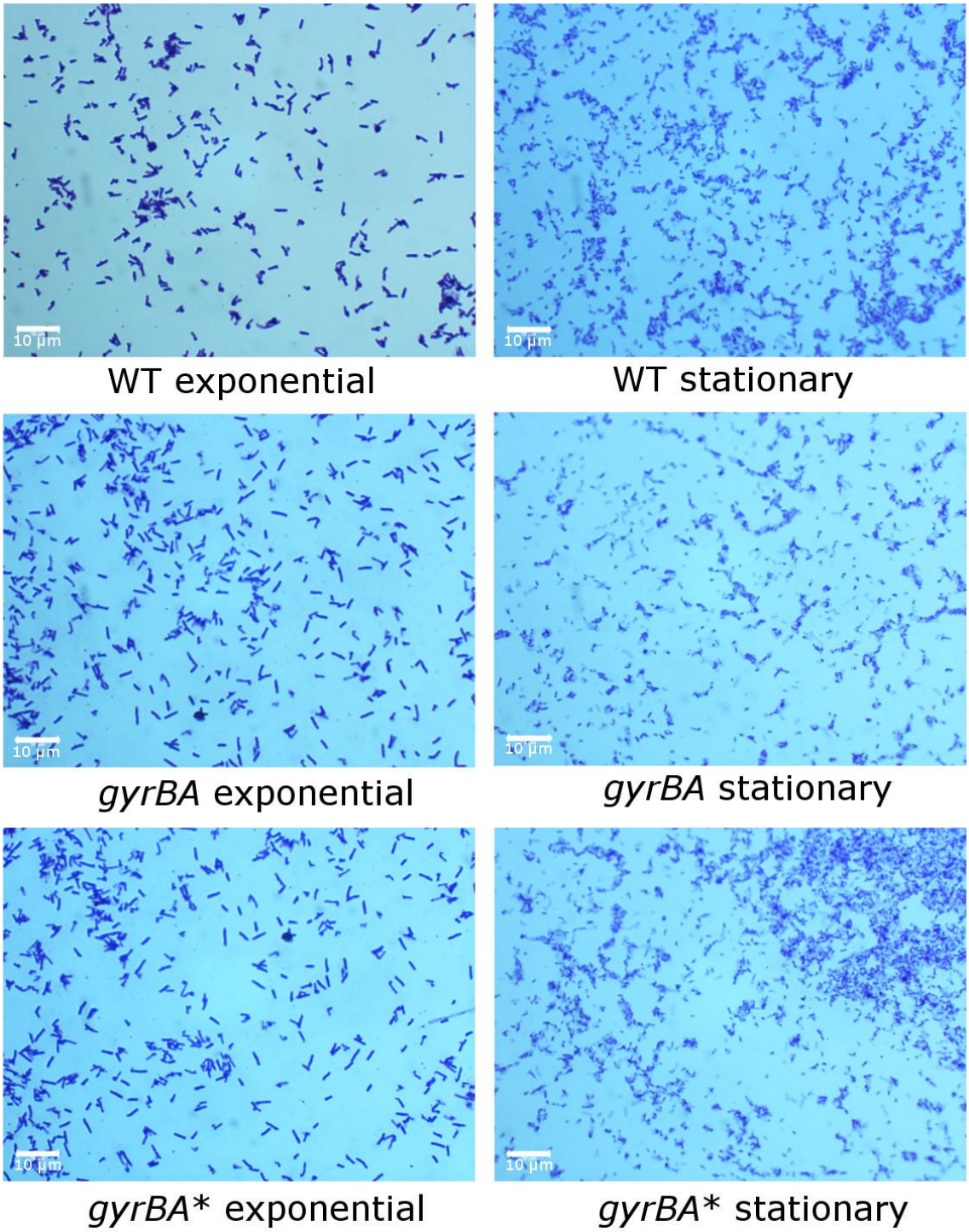


Figure 5.4 Morphology of the strains with the *gyrBA* operon. Bacteria were harvested, at the mid-exponential and at the late stationary growth phases, washed with PBS, heat-fixed, stained with crystal violet and viewed under 1000x magnification with oil immersion lens. Standard rod-shaped *Salmonella* cells were observed in the WT, the *gyrBA* and the *gyrBA** strains. All images are representative of three biological replicates. 10 µm bar is given for reference.

5.2.5 Competitive fitness alterations in *Salmonella* with the *gyrBA* operon

To test if expressing DNA gyrase from the operon altered the competitive fitness of *Salmonella*, derivatives of the WT and the *gyrBA* strains were made that carry a chloramphenicol resistance cassette within a pseudogene *SL1483* (Kröger *et al.*, 2013). This insertion has a neutral effect on fitness and allows us to use chloramphenicol to select for a strain that carries it. The competitions were carried out in the same flask, starting with equal numbers of cells of a "competitor" and a "WT". The WT was competed against the *gyrBA*+Cm^R and, as a control, the WT+Cm^R was competed against the *gyrBA* over the period of 24 h (Table 2.1). There was no difference in fitness detected between the WT and the *gyrBA* (Fig. 5.5).

Prior to working with the *gyrBA* strain, the *gyrBA** strain was competed with the WT in the same way. The *gyrBA** was found to be outcompeted by the WT (Fig. 5.5). However, this fitness deficiency is not an outcome of the operon construction, as the *gyrBA* strain has fitness that is equal to that of the WT. It is also not due to the growth defect observed in the *gyrBA** (Fig. 5.2a, Fig. 5.2b), as at 24 h growth of the *gyrBA** equalizes with growth of the WT. Instead, it is most likely to be another outcome of the SNP in *dxr*.

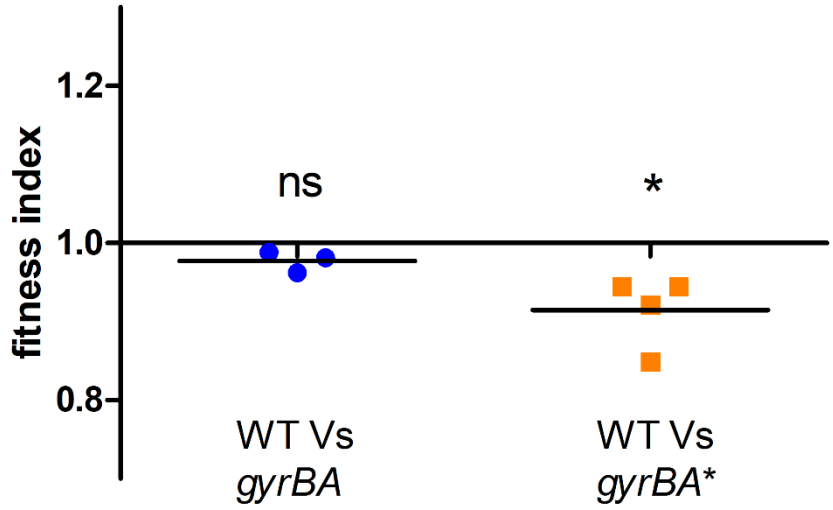


Figure 5.5 Competitive fitness of strains with the *gyrBA* operon. Fitness of a) the *gyrBA* b) the *gyrBA** compared to the WT SL1344 in LB broth grown for 24 h at 200 rpm at 37°C. Fitness index = 1 means that the competed strains are equally fit, f.i. < 1 indicates that the competitor strain is less fit than the WT, f.i. > 1 indicates that the competitor is more fit than the WT. The *GyrBA* and the WT are equally fit, while the *gyrBA** is significantly less fit than the WT. Significance was found by one sample T-test, $P<0.05$.

5.2.6 Effect of the *gyrBA* operon in *Salmonella* on cell motility

Although DNA gyrase is not associated with an effect on flagellar motility in *Salmonella*, motility of the *gyrBA* strain was measured alongside with strains that contain repositionings of IHF-encoding genes. A small but statistically significant decrease in motility of the *gyrBA* was found, but it is not likely to be biologically significant. The *gyrBA** strain displayed significant decrease in motility. An effect of the growth defect of the *gyrBA** strain on motility was eliminated by letting motility assay to proceed 20 min longer for the *gyrBA** than for the WT (Fig. 5.6). But again, as in the case with the competitive fitness decrease, the motility decrease was not an effect of the operon construction but was caused instead by the SNP in *dxr*.

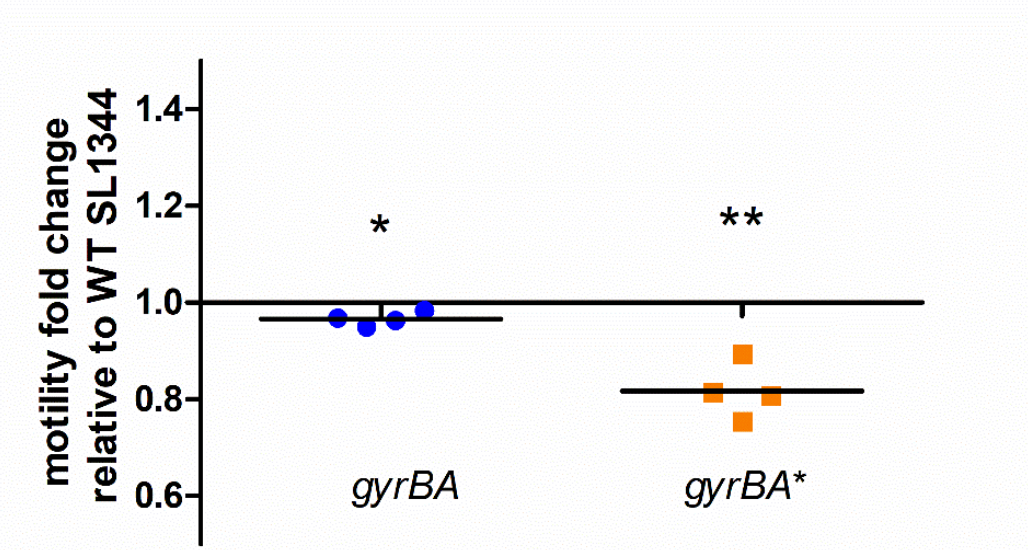


Figure 5.6 Motility of strains with the *gyrBA* operon relative to the WT. Diameters of swimming motility were measured after 5 h incubation at 37°C on soft 0.3% LB agar. The *gyrBA* is slightly but statistically significantly less motile than the WT. The *gyrBA** is significantly less motile than the WT. Values below 1 indicate that the strain is less motile than the WT. Significance was found by one sample T-test, where $P<0.05$.

5.2.7 Gene expression of *gyrA* and *gyrB* in the strain with the *gyrBA* operon is not altered

In *Salmonella gyrA* and *gyrB* – genes that encode DNA gyrase, are located at geographically separate loci (Fig. 1.1). This means that their expression is controlled independently of each other by different upstream and downstream regulatory regions. As a result of repositioning the *gyrA* ORF downstream of the *gyrB* ORF, *gyrA* becomes transcribed from the P_{gyrB} as part of a bicistronic mRNA. Dependency of *gyrA* expression on new regulatory inputs may affect its transcriptional patterns.

A gene dosage effect is one of the reasons that may affect *gyrA* expression when it is transcribed from the *gyrBA* operon. A gene that is close to *oriC* is present at a higher copy number than a gene close to Ter. This is because during rapid growth chromosome replication at *oriC* initiates before the previous full round of replication is completed. As a result, a gene located close to *oriC* may be present in a cell at higher copy number. This effect is particularly significant during exponential growth (Bryant *et al.*, 2014; Schmid & Roth, 1987). The *oriC*-proximal position is advantageous for genes that are required to be present at higher-copy number during rapid growth and beyond that, such as ribosomal and RNA polymerase genes (Soler-Bistué *et al.*, 2017), and could be similarly beneficial for the DNA gyrase-encoding genes. In *Salmonella* the *gyrB* gene is located in the NS-left macrodomain in the vicinity of the origin, while g*yrA* is in the Left macrodomain closer to the terminus (Fig. 1.1). In the *gyrBA* strain both genes are near the origin, increasing copy number of *gyrA*.

To compare the gene expression patterns of *gyrA* and *gyrB* between the *gyrBA* and the WT, RNA was extracted from both strains at various growth stages, DNase-treated and converted to cDNA. The abundance of *gyrA* and *gyrB* transcripts was determined by RT-qPCR using pairs of primers (SL_gyrA_qPCR_Pf, SL_gyrA_qPCR_Prev and SL_gyrB_F, SL_gyrB_R, Table 2.2). Levels of the transcripts were normalized against *hemX* – a gene whose expression does not change under tested conditions (Kröger *et al.*, 2013). Primers (hemX_qPCR_Pf and hemX_qPCR_Prev, Table 2.2) were used to determine the abundance of *hemX* transcript.

The levels of both the *gyrA* and *gyrB* transcripts were decreasing from the early exponential to the stationary phase and were present in approximately a 1:1 ratio (Fig. 5.7), as shown previously (Cameron *et al.*, 2011). Such gene expression pattern is not surprising since the *gyrA* and *gyrB* encode subunits of the same enzyme that must be present in equal amounts to assemble into a functional heterotetramer. It is surprising, however, that the *gyrB* and *gyrA* genes are transcribed at the same level in the WT, despite gene dosage differences, and may indicate the presence of a strict post-transcriptional regulation that brings *gyrA* and *gyrB* mRNA to the same level. No change in the expression of either gene was found in the *gyrBA* strain when compared to the WT. This suggests that in spite of being arranged into the operon and producing hybrid *gyrBA* mRNA species, *gyrB* and *gyrA* gene expression remains the same. This could be because *gyrA* is positioned downstream of the gene that has similar or identical regulation, so that *gyrA* is transcribed in the same way as at its original locus.

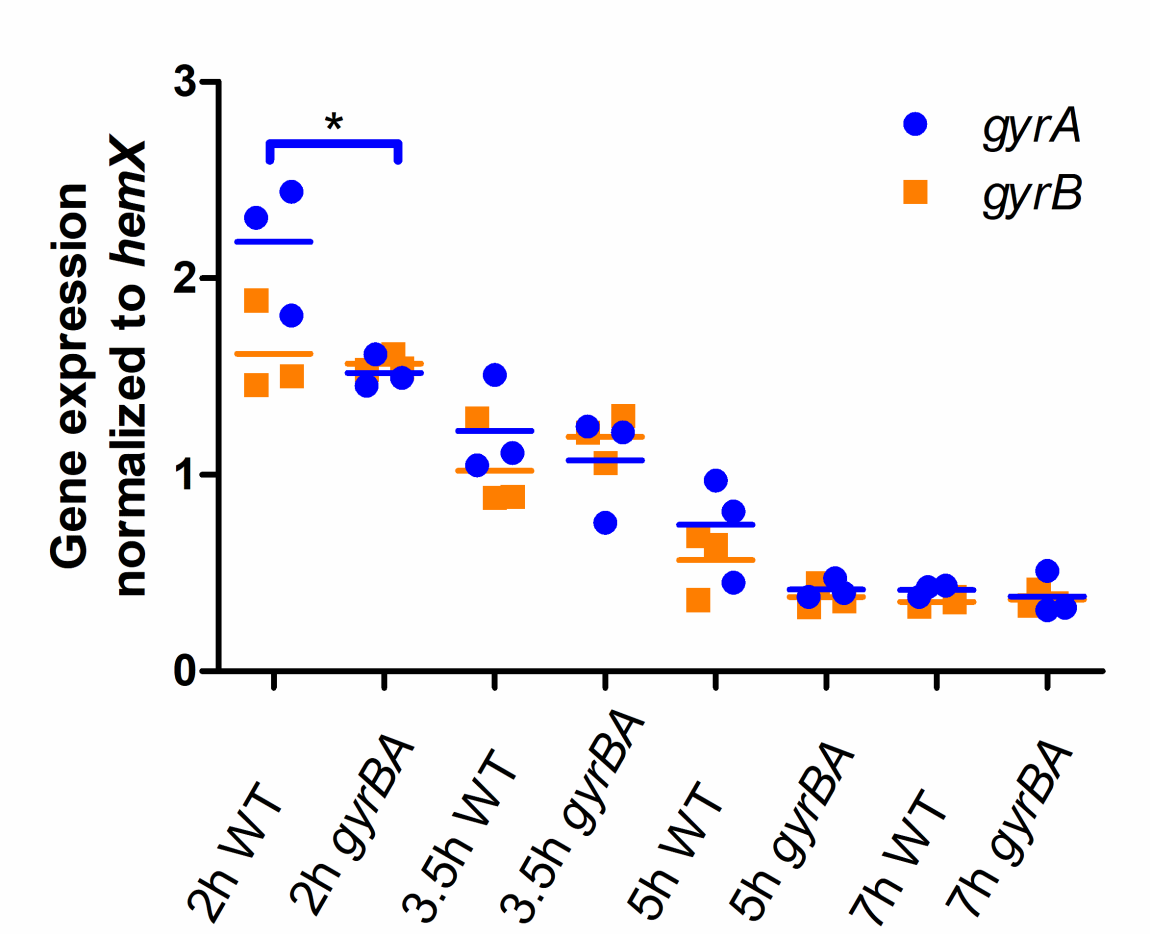


Figure 5.7 Gene expression of *gyrA* and *gyrB* in the WT and the *gyrBA*. Cells were grown in LB broth at 37°C at 200 rpm and samples were taken at 2 h, 3.5 h, 5 h and 7 h representing lag, exponential, exponential-stationary transition and early stationary phases of growth, respectively. Gene expression of *GyrA* and *gyrB* decreased from 2 h to 7 h. *GyrA* gene expression was lower in the *gyrBA* strain than in the WT at 2 h and comparable at other time points. Three biological replicates were done. Significance was found by unpaired Student's T-test, where $P<0.05$.

5.2.8 Minimal inhibitory concentrations of antibiotics that act on DNA gyrase: values for the WT and the strain with the *gyrBA* operon

DNA gyrase is essential for bacteria, making it an attractive target for antibiotic development. Two main classes of antibiotics targeting different subunits of DNA gyrase exist. Antibiotics in the aminocoumarin class compete with ATP for binding to the GyrB subunit of DNA gyrase (Lewis *et al.*, 1996), leading to tight binding of the aminocoumarin to GyrB and inhibition of the energy-generating ATPase activity of DNA gyrase. Another class – quinolones – bind to the GyrA subunit of DNA gyrase, trapping gyrase-DNA complexes and leading to accumulation of double stranded DNA breaks (Drlica & Zhao, 1997). The possibility that the introduction of the *gyrBA* operon might have changed the susceptibility of *Salmonella* to these antibiotics was investigated.

To exploit this, minimal inhibitory concentrations of aminocoumarins – novobiocin and coumermycin; and quinolones – nalidixic acid and ciprofloxacin were measured in the WT and the *gyrBA* strains. The bacterial strains were grown in a 96-well plate with 1:1.25 serially diluted antibiotics in LB broth for 18 h at 37°C at 200 rpm. OD_{600} was measured and MIC_{90} was defined as antibiotic concentration at which 90% of the bacterial growth was inhibited. Only partial antibiotic dilution ranges were plotted, to show where a change in cell density was observed (Fig. 5.8).

Statistically significant differences in growth between the WT and the *gyrBA* strains were seen in the presence of aminocoumarins (Fig. 5.8a and Fig. 5.8b). While the growth of the WT and the *gyrBA* in the presence of quinolones was identical (Fig. 5.8c and Fig. 5.8d). Interestingly, opposing response to the different drugs in the aminocoumarin class was observed: the *gyrBA* strain was more resistant than the WT to novobiocin, displaying higher MIC_{90} and higher cell densities at sub-inhibitory concentrations; at the same time, the *gyrBA* strain was less resistant to coumermycin, with the effect seen at sub-inhibitory concentrations.

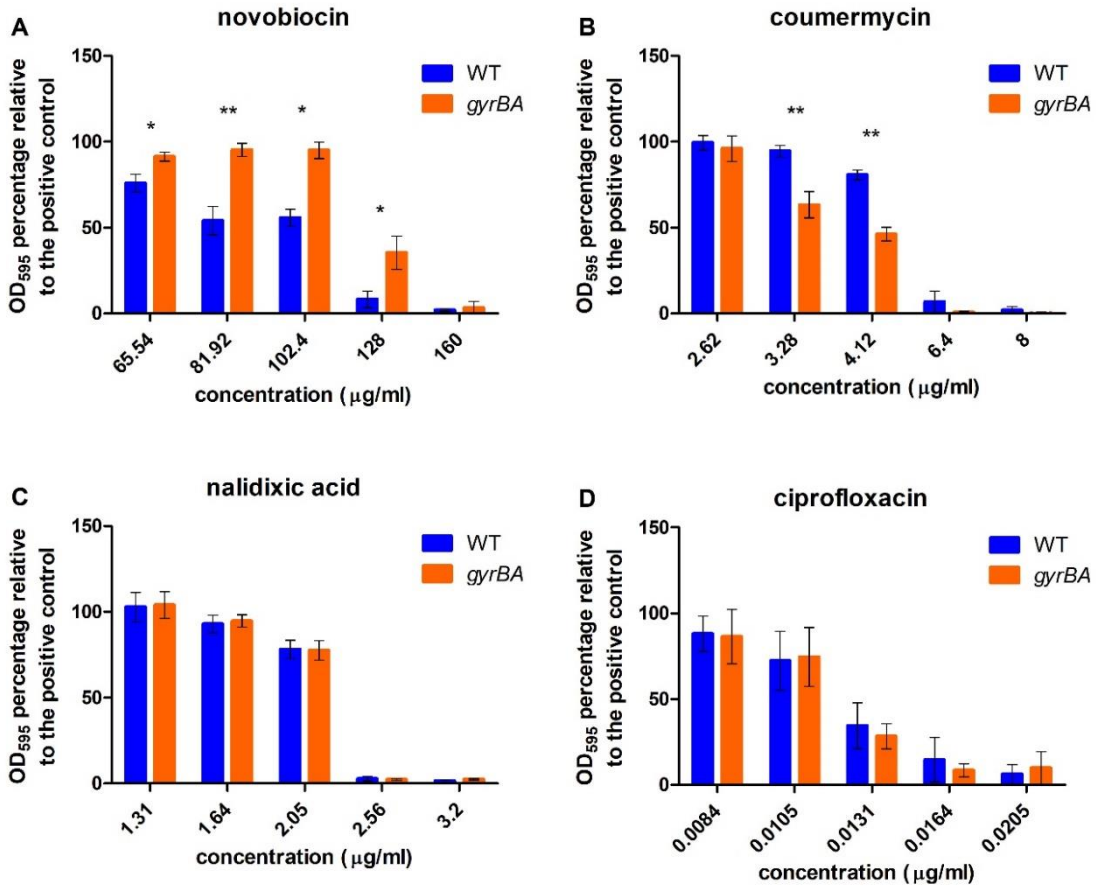


Figure 5.8 MICs of DNA gyrase-targeting drugs in the SL1344 and the *gyrBA* strains. Cells were grown in a 96-well plate with 1:1.25 serially diluted antibiotics in LB broth for 18 h at 37°C and 200 rpm. Cell density by OD_{600} was measured. a) Percentage survival of the WT and the *gyrBA* in 65.54-160 μg/ml novobiocin. MIC_{90} of the WT = 128 μg/ml, MIC_{90} of the *gyrBA* = 160 μg/ml. b) Percentage survival of the WT and the *gyrBA* in 2.62-8 μg/ml coumermycin, MIC_{90} = 6.4 μg/ml. c) Percentage survival of the WT and the *gyrBA* in 1.31 – 3.2 μg/ml nalidixic acid, MIC_{90} = 2.56 μg/ml. d) Percentage survival of the WT and the *gyrBA* in 0.0084 – 0.0205 μg/ml ciprofloxacin, MIC_{90} = 0.0164 μg/ml. Error bars represent standard deviation between at least three biological replicates. Significance was found by unpaired Student's T-test, where $P<0.05$.

5.2.9 SPI-1/2 gene expression in the *gyrBA* operon and WT strains

Salmonella pathogenicity islands (SPIs) are horizontally acquired chromosomal regions that each code for a type III secretion system that *Salmonella* uses to enter a host cell (SPI-1) and to escape it after replication (SPI-2) (Fig. 1.7) (Cirillo *et al.*, 1998; Galán, 1996; Shea *et al.*, 1996). DNA gyrase, together with topoisomerase I, ensure balanced organisation of the global DNA supercoiling through a homeostatic regulatory mechanism (DiNardo *et al.*, 1982; Pruss *et al.*, 1982). SPI-1 and SPI-2 gene promoters are activated by OmpR binding that occurs during DNA relaxation (Cameron & Dorman, 2012), indicating that these promoters are sensitive to alterations in DNA supercoiling. The repositioning of *gyrA* may have led to an altered ability of DNA gyrase to supercoil DNA, affecting the expression of SPI-1 and/or SPI-2. Another reason for SPI expression measurements is that it is a useful indicator of whether *Salmonella* is able to establish a successful infection. The *gfp* chromosomal reporter fusions to the promoters of archetypal genes *prgH* (SPI-1) and of *ssaG* (SPI-2) were transduced into the WT and the *gyrBA* strains (Table 2.1) and used to measure SPI-1 and SPI-2 expression activity (Ibarra *et al.*, 2010).

GFP fluorescence at 528 nm was measured every 20 min over 24 h and divided by OD_{600} for normalization. In the *gyrBA* strain both SPI-1 and SPI-2 expression patterns were identical to those in the WT when grown in LB broth (Fig. 5.9). This indicates that DNA gyrase expressed from the *gyrBA* operon in *Salmonella* does not affect expression of SPIs. However, growth in rich medium, such as LB broth, does not reflect what *Salmonella* experiences in its environment.

Low magnesium treatment is a signal that was exploited in this work to mimic environment inside a macrophage. To vary the magnesium ions concentration, a defined N-minimal medium was used (Nelson & Kennedy, 1971) that contained low (10 μM) or high (10 mM) concentration of Mg^{2+} ions.

SPI-1 and SPI-2 expression was measured in a minimal medium supplemented with Mg^{2+} at both high and low concentrations, culturing bacteria as in 2.2.2. SPI expression is a costly process (Sturm *et al.*, 2011), thus, *Salmonella* can be expected to shut SPI expression during growth in a minimal medium. Indeed,

SPI-1 expression was turned off in a minimal medium at both Mg^{2+} concentrations (Fig. 5.10a and Fig. 5.10c). SPI-2 expression was similarly turned off in a minimal medium at high Mg^{2+} concentration (Fig. 5.10b), however, it was on at low Mg^{2+} concentration (Fig. 5.10d). At this condition SPI-2 expression was decreased in the *gyrBA* strain at the stationary phase of growth. This indicates that DNA gyrase expressed from the *gyrBA* operon in *Salmonella* affects expression of SPI-2 in condition that resembles environment inside the *Salmonella*-containing vacuole.

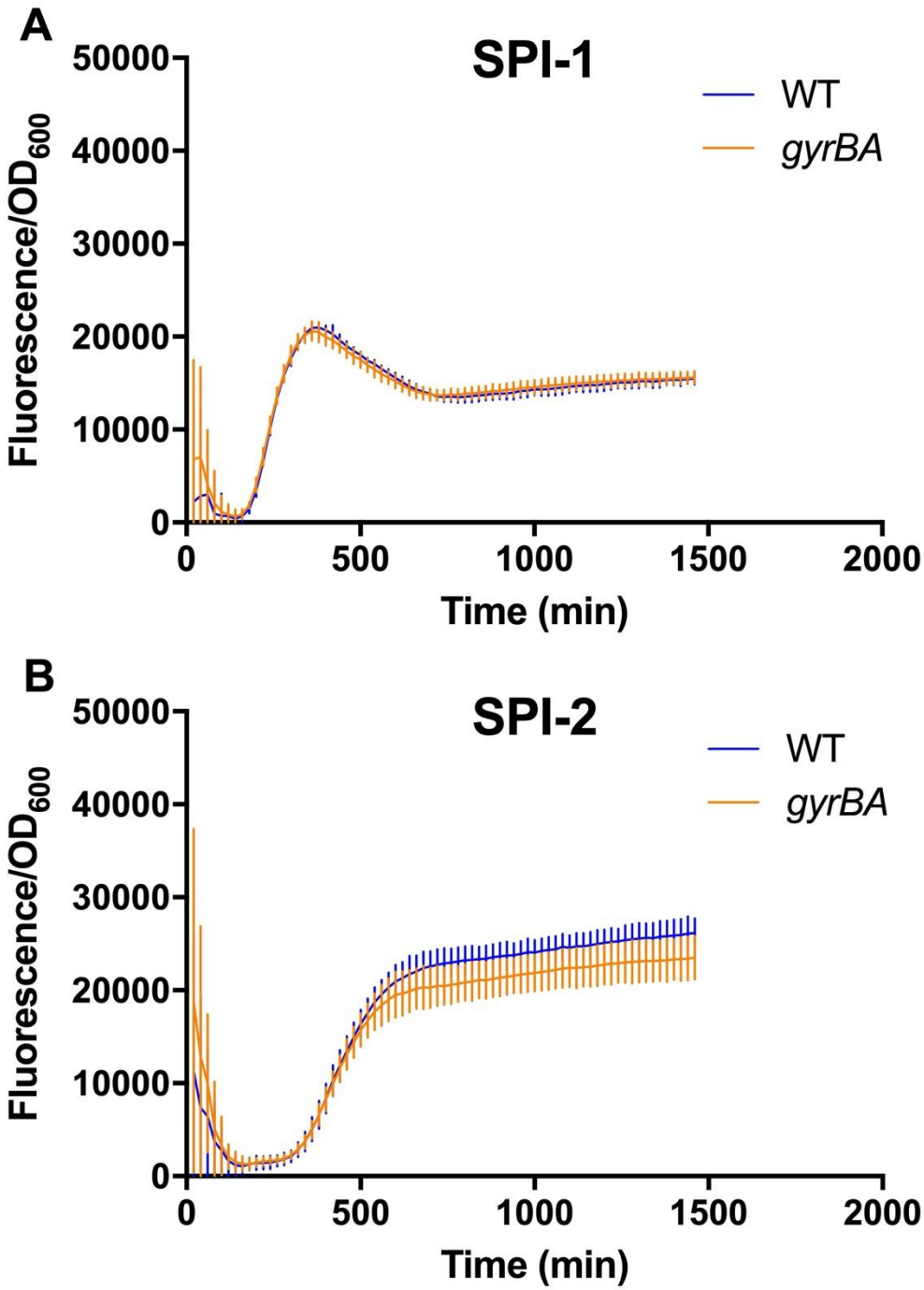


Figure 5.9 *Salmonella* pathogenicity islands expression measured with aid of *gfp* reporter fusions in the *gyrBA* strain compared to the WT grown in LB broth. Fluorescence at 528 nm normalized against OD_{600} was plotted for every 20 min measurement over 24 h. a) SPI-1 expression in the *gyrBA* was identical to that in the WT. b) SPI-2 expression in the *gyrBA* was identical to that of the WT. All plots are results of at least three biological replicates, error bars represent standard deviation.

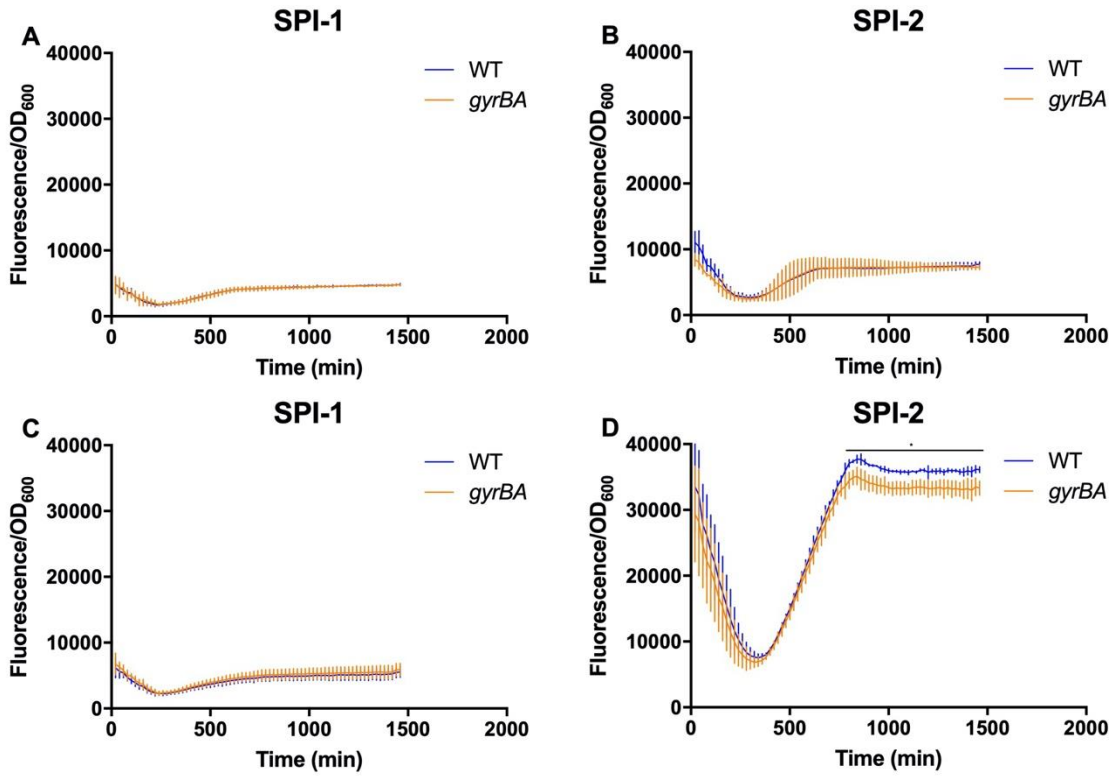


Figure 5.10 *Salmonella* pathogenicity islands expression measured with aid of *gfp* reporter fusions in the *gyrBA* strain compared to the WT grown in a minimal medium with high and low magnesium concentrations. Fluorescence at 528 nm normalized against OD_{600} was plotted for every 20 min measurement over 24 h. a) SPI-1 expression in a minimal medium with high Mg^{2+} concentration was off in both the WT and the *gyrBA*. b) SPI-2 expression in a minimal medium with high Mg^{2+} concentration was off in both the WT and the *gyrBA*. c) SPI-1 expression in a minimal medium with low Mg^{2+} concentration was off in both the WT and the *gyrBA*. d) SPI-2 expression in a minimal medium with low Mg^{2+} concentration was lower in the *gyrBA* strain than in the WT at the stationary phase of growth. All plots are results of at least three biological replicates, error bars represent standard deviation. Asterisk represents significance for timepoints highlighted with a black curve. Significance was found by Student's unpaired T-test, where $P<0.05$.

5.2.10 Global DNA supercoiling is altered in the *gyrBA* strain

The primary and unique function of DNA gyrase among other topoisomerases is to introduce negative supercoils into DNA. Negative DNA supercoiling introduced by DNA gyrase results in the overall negatively supercoiled chromosome and is used to compact DNA at multiple levels of the nucleoid's organisation (Dorman, 2013). It was anticipated that repositioning genes that encode DNA gyrase may influence DNA supercoiling set points in a cell due to a change in gyrase productivity or a change in the way it senses intracellular environment.

Measuring the supercoiling of the entire bacterial chromosome is a technically-demanding task that involves determining the sedimentation coefficient of the isolated, intact chromosome in a sucrose gradient or the rate of incorporation into the chromosome of the UV-sensitive trimethylpsoralen dye (Drlica & Snyder, 1978; Lal *et al.*, 2016). For routine purposes, reporter plasmid electrophoresis in the presence of a DNA intercalating dye is a more practical option (Pruss *et al.*, 1982). The small, high-copy number plasmid pUC18 was transformed into the WT and the *gyrBA* strains and used to monitor *Salmonella's* global DNA supercoiling at the stationary phase of growth in cultures grown in LB broth. Plasmid DNA samples of pUC18 were extracted from the WT and the *gyrBA* strains, as in 2.5.2, and electrophoresed on a chloroquine gel, as in 2.23.1.

At the concentration of chloroquine used here, the more negatively supercoiled topoisomers run further to the right in the sample lane, while the more relaxed topoisomers are retained and appear to the left of the sample lane (Fig. 5.11a). Reporter plasmid DNA in the *gyrBA* strain was found to be more relaxed, showing a linking number change (ΔLk) of +1 compared to the WT. Each lane was scanned densitometrically and the intensity profiles created in ImageJ are shown on top of the sample lanes for simple visual interpretation (Fig. 5.11a). Each band represents a single topoisomer of pUC18. While Fig. 5.11a highlights the supercoiling difference, Fig. 5.11b shows a section of the whole chloroquine gel. In addition to monomers, plasmid pUC18 tends to form dimers and higher order oligomers and these are displayed in Fig. 5.11b.

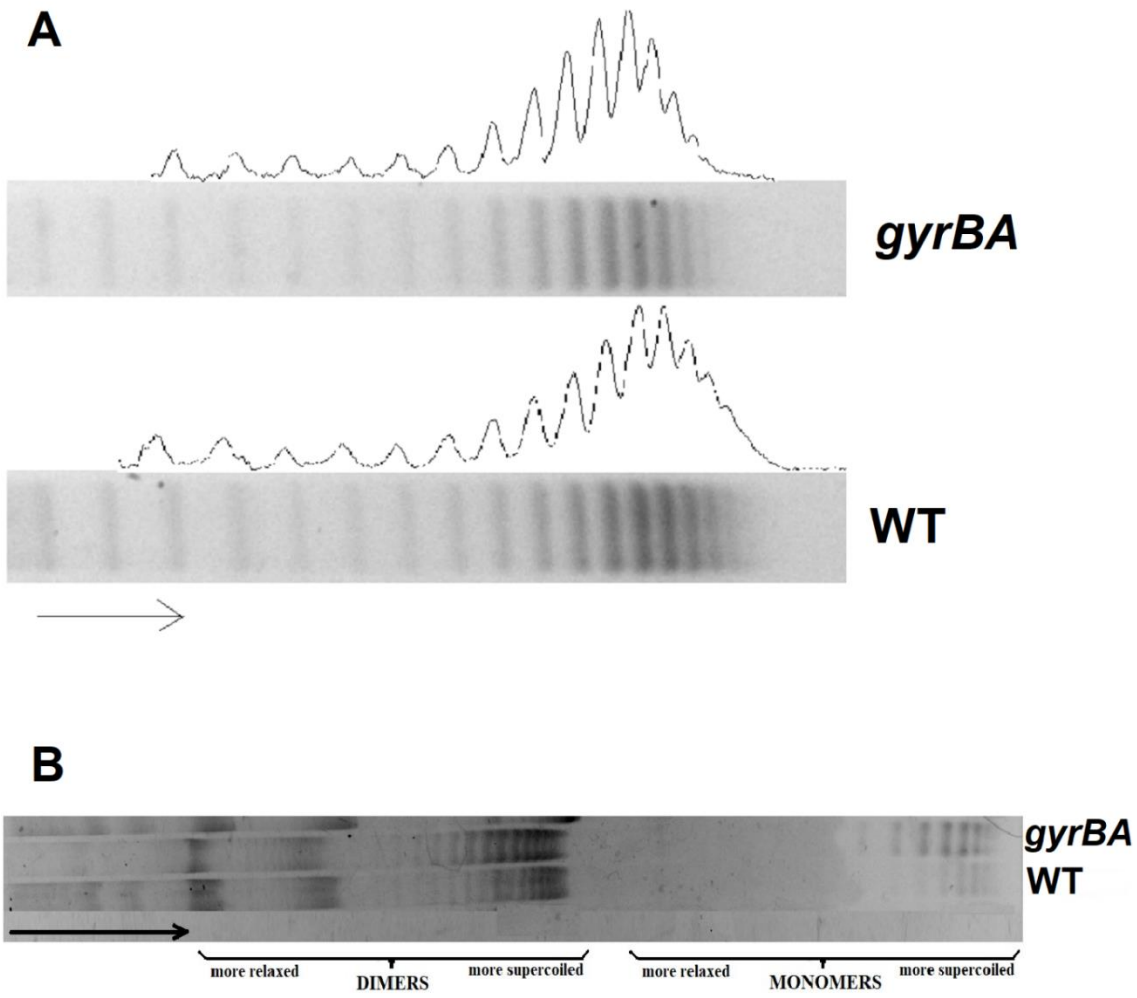


Figure 5.11 Global DNA supercoiling in the SL1344 WT and the *gyrBA* strains grown in LB broth. Samples of pUC18 plasmid were extracted from the WT and the *gyrBA* at the stationary growth phase and run on a 0.8% agarose gel containing 2.5 µg/ml chloroquine to separate pUC18 topoisomers according to the degree of DNA supercoiling. Arrow shows the direction of electrophoretic flow with the more supercoiled plasmid topoisomers at the right of the gel. a) Global DNA supercoiling pattern of the WT and the *gyrBA* when grown in LB. Sample lanes are supplemented with intensity profiles. b) An example of a full chloroquine gel. Analysis is representative of four biological replicates. Densitometry profiles were generated with ImageJ.

5.2.11 Global DNA supercoiling in the *gyrBA* strain growing in a minimal medium with variable Mg^{2+} concentration

The macrophage is an important immune cell that senses and engulfs pathogens, exposing them to a hostile environment in a vacuole that includes low pH, reactive oxygen and nitrogen species, low concentration of magnesium ions, etc. While such conditions are designed to kill bacteria, *Salmonella* has evolved the capacity to survive and proliferate in the *Salmonella*-containing vacuole (SCV) of mammalian macrophages making the macrophage an important niche for *Salmonella* during infection (Ilyas *et al.*, 2017). Some of these intramacrophagal factors are important signals for *Salmonella* to express its virulence genes (Fass & Groisman, 2009; Srikumar *et al.*, 2015).

Similar to the SPI expression study, low magnesium treatment was exploited in the global DNA supercoiling determination study to mimic environment inside a macrophage. To vary the magnesium ions concentration, a defined N-minimal medium was used (Nelson & Kennedy, 1971) that contained low (10 μM) or high (10 mM) concentration of Mg^{2+} ions.

Reporter plasmid DNA in the *gyrBA* strain was found to be more relaxed than in the WT when cultured in LB broth (Fig. 5.11). Interestingly, when global DNA supercoiling was assessed in N-minimal medium, the pattern was the opposite to that seen in LB: DNA in the *gyrBA* was more supercoiled than DNA in the WT (Fig. 5.12a). In a medium with a high concentration of Mg^{2+} (10 mM) Δ*Lk* was -2, while at the low concentration of Mg^{2+} (10 μM) the *gyrBA* strain displayed the smaller difference from the WT (Δ*Lk* = -1). General relaxation of about Δ*Lk* = +3 was equally seen in both the WT and the *gyrBA* in low Mg^{2+} compared to the high Mg^{2+}, and was in line with previously published data that *Salmonella*'s DNA relaxes inside a macrophage (O'Croinin *et al.*, 2006). This also indicates that *in vitro* relaxation can be achieved not only in response to a change in pH (Colgan *et al.*, 2018) but also by a change in a single chemical parameter, such as Mg^{2+} concentration. While Fig. 5.12a highlights the DNA supercoiling difference, Fig. 5.12b shows a section of the whole chloroquine gel.

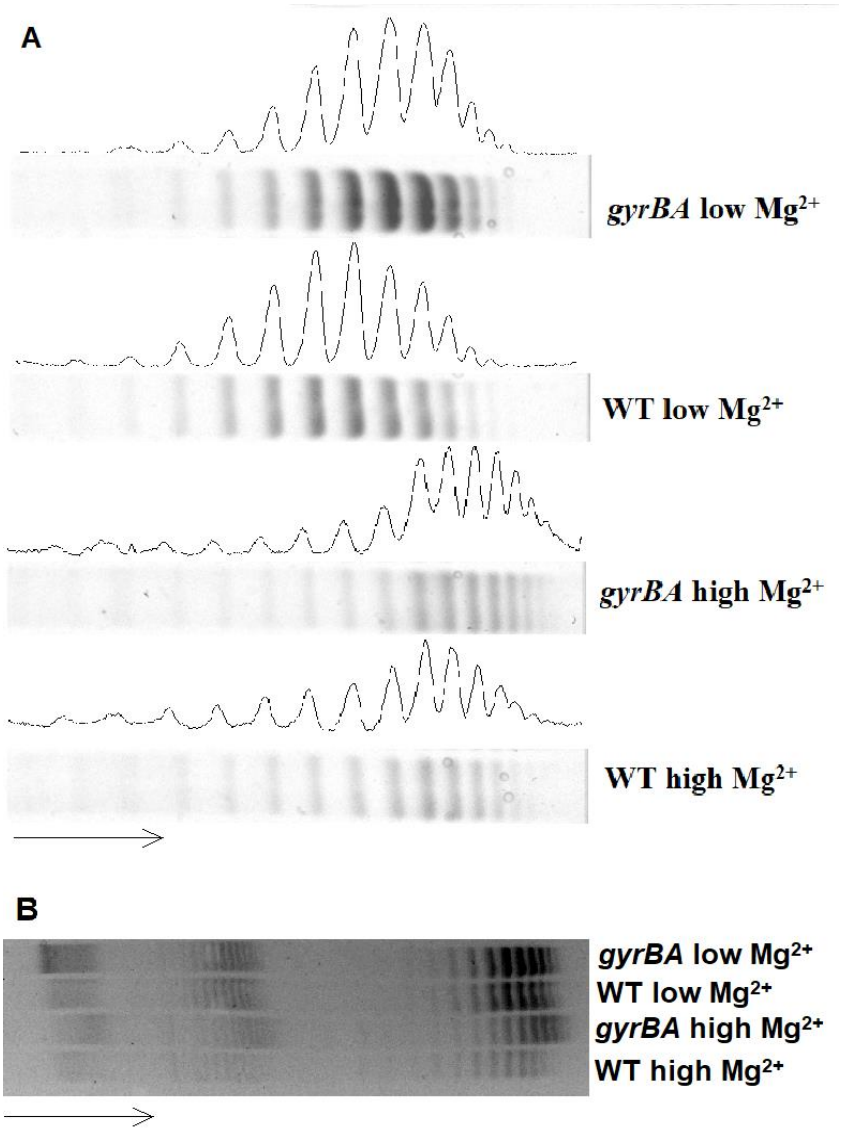


Figure 5.12 Global DNA supercoiling in the SL1344 WT and the *gyrBA* strains, grown in minimal medium with high and low magnesium concentrations. Samples of pUC18 plasmid were extracted from the WT and the *gyrBA* at the stationary growth phase and run on a 0.8% agarose gel containing 2.5 µg/ml chloroquine to separate pUC18 topoisomers according to the degree of DNA supercoiling. Arrow shows the direction of electrophoretic flow with the more supercoiled plasmids at the right of each sample lane. a) Global DNA supercoiling pattern of the WT and the *gyrBA* when grown in N-minimal medium with high (10 mM) Mg^{2+} and low (10 µM) Mg^{2+}. Sample lanes are supplemented with intensity profiles. b) An example of a full chloroquine gel. Analysis is representative of four biological replicates. Densitometry profiles were generated with ImageJ.

5.2.12 Ability of the *gyrBA* strain to infect and survive in murine macrophages

The decrease in SPI-2 expression during growth in a minimal medium with low Mg^{2+} concentration (Fig. 5.10d) suggests that *Salmonella* that expresses the *gyrBA* operon may not be able to survive within macrophages as successfully as the WT. This is because the full-scale expression of the *Salmonella* pathogenicity islands is required to provide bacteria with all necessary machinery to infect epithelial cells and strive within the *Salmonella*-containing vacuole of the mammalian macrophage – *Salmonella's* major niche during infection (Ilyas *et al.*, 2017).

In addition, global DNA supercoiling measurements showed that *Salmonella* expressing the *gyrBA* operon supercoils DNA to a different extent in response to conditions that mimic aspects of the environment inside the mammalian macrophage vacuole (in terms of Mg^{2+} concentration) when compared to the WT (Fig. 5.12). Since DNA supercoiling is a known global regulator of gene expression (Dorman & Dorman, 2016), potential alterations in DNA supercoiling of the *gyrBA* strain within a macrophage vacuole may cause alterations of gene expression in this strain. These expression changes could be significant enough to influence survival and replication of the *gyrBA* strain inside a *Salmonella*-containing vacuole (SCV).

Salmonella can enter a macrophage via two modes: (1) SPI-1 mediated active invasion and (2) passive host-mediated phagocytosis (Drecktrah *et al.*, 2006). Both modes are infection-relevant and were used to set up infections of RAW264.7 murine macrophages. To induce SPI-1 mediated entry, bacteria were grown to mid-exponential phase before infection at MOI 5 (Steele-Mortimer *et al.*, 1999). To promote phagocytosis, bacteria were grown to late stationary phase and complement-opsonised by incubation with normal mouse serum before infection at MOI 20 (Lee & Falkow, 1990). The infection mix was enumerated, and gentamycin protection assays were carried out, as in 2.22. This allowed assessment of the performance of the bacterium at different stages of the infection process: efficiency of entry to the macrophage, intracellular survival and replication over a 20-h period post-infection.

Salmonella cells inside macrophages were enumerated by viable counts at 1 h, 3 h, 8 h, 16 h and 20 h post-infection. 1 h was denoted as the reference point for entry and all other timepoints were normalized against it.

As expected, no difference in either entry by phagocytosis (Fig. 5.13a) or survival and replication (Fig. 5.13b) was observed between the *gyrBA* and the WT strain. However, differences were found when infections were carried out using the SPI-1 entry model.

Surprisingly, an increase in SPI-1 mediated entry efficiency was observed in the *gyrBA* (Fig. 5.14a). This is despite the lack of observable changes in SPI-1 expression (Fig. 5.9 and Fig. 5.10). The environment inside the SCV is comprised of multiple factors, while the Mg^{2+} ions concentration was the only SCV-associated stimulus applied in the experimentation with the N-minimal medium. Usage of alternative growth conditions that specifically cause SPI-1 induction might discern the potential SPI-1 expression differences between the WT and the *gyrBA* strains.

When the SPI-1 mediated entry model was utilized, a killing event was observed during the 3 h and 8 h p.i. in both strains. Later, at 16 h p.i., *Salmonella* appeared to adapt and replicate significantly, followed by a further decline by 20 h p.i., possibly caused by macrophage cell death (van der Velden *et al.*, 2000). The survival and replication of the *gyrBA* strain in a murine macrophage was decreased significantly at 16 h post-infection (Fig. 5.14b). This is justified by the decrease in SPI-2 expression that was observed in a minimal medium with low Mg^{2+} (Fig. 5.10d). Overall, these results demonstrated that DNA gyrase expressed from the *gyrBA* operon in *Salmonella* improved the ability of bacterium to enter via SPI-1 mediated mechanism but impaired the ability to survive within murine macrophages.

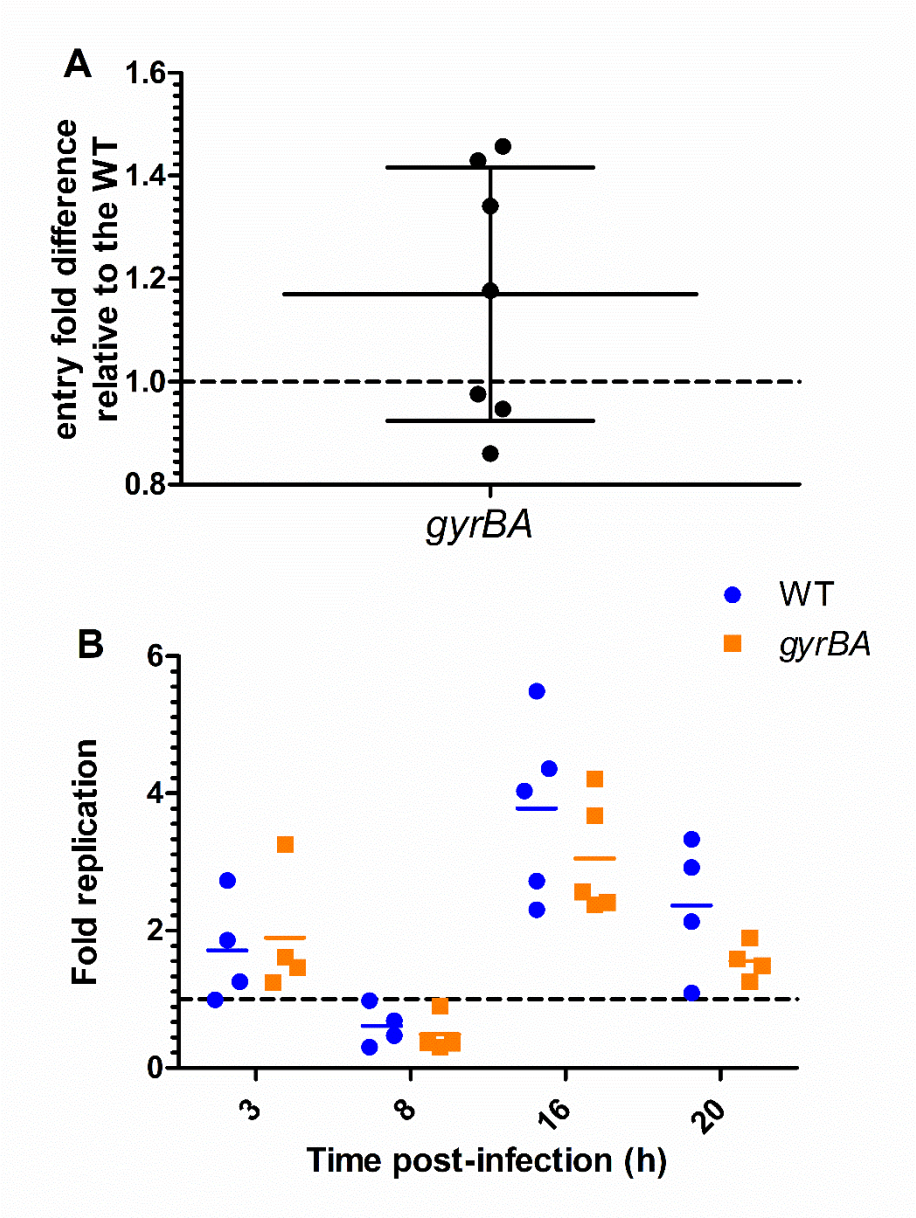


Figure 5.13 Entry by phagocytosis and survival of the WT and the *gyrBA* in RAW264.7 macrophages. Cells were infected with SPI-2 induced bacteria, grown to stationary phase and complement-opsonised to promote phagocytosis. a) Entry was measured by enumerating CFU 1 h post-infection. Each replicate was normalized to the infection mix and the WT. Mean and standard deviation are shown. No significant difference was found by one sample T-test, where $P<0.05$. b) Survival and replication were measured by enumerating CFUs at 3 h, 8 h, 16 h and 20 h post-infection. Fold replication represents the CFU recovered at a particular time point divided by the CFU at 1 h. Mean and individual replicates are shown. No significant difference was found by unpaired Student's T-test.

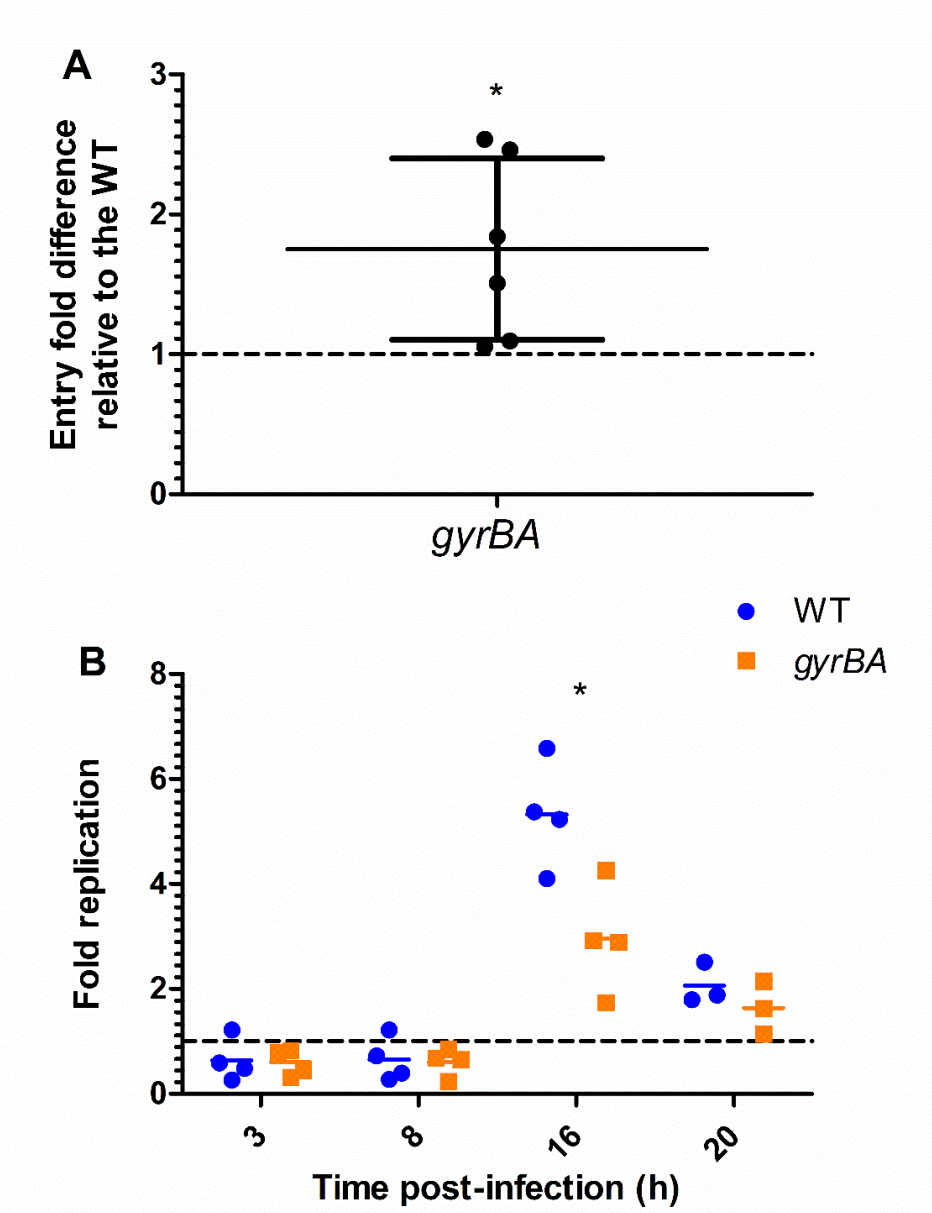


Figure 5.14 SPI-1 mediated entry and survival of the WT and the *gyrBA* in RAW264.7 macrophages. Cells were infected with SPI-1 induced bacteria, grown to mid-exponential phase and to promote active SPI-1 mediated invasion. a) Entry was measured by enumerating CFU 1 h post-infection. Each replicate was normalized to the infection mix and the WT. Mean and standard deviation are shown. Significance was found by one sample T-test, where $P<0.05$. b) Survival and replication were measured by enumerating CFUs at 3 h, 8 h, 16 h and 20 h post-infection. Fold replication represents the CFU recovered at a particular time point divided by the CFU at 1 h. Mean and individual replicates are shown. Significance at 16 h was found by unpaired Student's T-test, where $P<0.05$.

5.2.13 Locations of the *gyrA* and *gyrB* genes in bacteria

In *Salmonella* and *Escherichia coli* – members of the same family (the *Enterobacteriaceae*), the positions of *gyrA* and *gyrB* – the genes that encode DNA gyrase – are conserved (Sobetzko *et al.*, 2012). The *gyrA* gene is located in the Left and the *gyrB* in the NS-left macrodomains (Fig. 1.1). However, in other bacteria the positions of these genes differ. For example, *gyrBA* operons are found in *Mycobacterium* spp, *Borrelia* spp, *Streptomyces* spp and others (Szafran *et al.*, 2016; Unniraman & Nagaraja, 1999). To gain a deeper understanding of the relative positions of *gyrA* and *gyrB* across bacteria, the positions of the DNA gyrase-encoding genes were determined in multiple species.

The DoriC 10.0 online database (available to download at tubic.org/doric) lists locations of origin sequences in multiple bacterial and archaeal genomes (Luo & Gao, 2019). Locations of *gyrA* or *gyrB* were visualised on a chromosome of a particular species using Ensembl bacteria browser (bacteria.ensembl.org). DoriC was used to find the location of *oriC* on the chromosome of the species under investigation. Distance in base pairs between the *oriC* and the gene was calculated and converted into the percentage of the total chromosome size. An attempt was made to cover bacterial taxonomy as broadly as possible, encompassing members of the major bacterial phyla, well-studied, and clinically important organisms in the analysis (Table 5.1). The table is neither complete nor claiming to include all the existing possibilities of *gyrA* and *gyrB* arrangements in bacterial chromosomes, but instead, exemplifies the arrangement possibilities mentioned in this work. Closely related species and those belonging to the less diverse phyla were found to share the chromosomal positions of *gyrA* and *gyrB* frequently. Thus, one representative of a taxonomic rank was often deemed sufficient for the purpose of inclusion into the table. Lower classification ranks were analysed within more diverse and studied phyla.

Relative inversions between the left and the right replicores were frequently seen and had no visible pattern. This is in agreement with the previous finding that while distance to the origin is highly conserved, inversions of genes around the Ter region of a chromosome are frequent and well-tolerated between *E. coli* and *Salmonella* (Alokam *et al.*, 2002). Various relative arrangements of *gyrA* and *gyrB* were observed: group 1 – *gyrB* and *gyrA* positioned separately, *gyrB*

near *oriC*, 5' of *gyrB* conserved; group 2 – *gyrB* and *gyrA* positioned separately, variable position of *gyrB*; group 3 – the *gyrBA* operon in the immediate vicinity of *oriC*, 5' of *gyrB* conserved; group 4 – the *gyrBA* operon away from *oriC*, variable position of *gyrB*. The arrangements of *gyrA* and *gyrB* genes were categorised into the four types not only according to the relative positions of *gyrA* and *gyrB* but mainly according to a parameter that was deemed more important – the degree of conservation of the genetic environment 5' of *gyrB*.

Table 5.1 suggests that all members of the class gammaproteobacteria (phylum Proteobacteria), including *E. coli* and *Salmonella*, some alphaproteobacteria, betaproteobacteria and epsilonproteobacteria were group 1. Some alpha, beta and deltaproteobacteria; members of the family *Streptococcaceae* (order Lactobacillales, phylum Firmicutes); members of class Flavobacteriia; multiple members of the phylum Bacteroidetes; Acidobacteria and *Deinococcus radiodurans* were group 2. Members of the phylum Actinobacteria, classes Clostridia and Bacilli (phylum Firmicutes), family *Enterococcaceae* and family *Lactobacillaceae* (order Lactobacillales, phylum Firmicutes) order Fusobacteria (phylum Fusobacteria) and phylum Terenicutes were group 3. Finally, members of the phyla Chlamydiae were group 4. There is perhaps more variation within the group 4, but it is considerably smaller and was not found by this method. *Mycoplasma* is an anomaly of the group 3, since not all its species clearly belong to this group. Some *Mycoplasma* possess the conserved genes 5' of *gyrB* close enough but not in the immediate vicinity. However, the orientation of genes 5' of *gyrB* remains favourable for the initiation of its transcription (see discussion to this chapter), therefore, I placed *Mycoplasma* into the group 3. From what was analysed, members of the same taxonomic rank do not necessarily have to belong to the same group, especially in diverse phyla. For example, both group 2 and group 3 arrangements are present within the Firmicutes. Moreover, both arrangements are present within the order Lactobacillales alone. Some less diverse phyla such as Fusobacteria and Chlamydiae belong to only one group. No variation within families was found in all the cases.

Organism (phylum, lowest clade sharing the arrangement)	***gyrB* to *oriC* b.p./replicore**	***gyrB* to *oriC* % of total chromosome**	***gyrA* to *oriC* b.p./replicore**	***gyrA* to *oriC* % of total chromosome**
Group 1				
Escherichia coli	45409 right	0.98%	1585679 right	34.20%
Salmonella enterica serovar Typhimurium	42033 left	0.86%	1730509 left	35.48%
Salmonella enterica serovar Gallinarum	37634 right	0.81%	1382662 left	29.68%
Shigella flexneri	57368 left	1.25%	1577518 left	34.24%
Yersinia pestis	37176 left	0.79%	1073028 right	22.82%
Vibrio cholerae	10386 right	0.35%	1329436 right	44.90%
Pseudomonas aeruginosa	2248 right	0.04%	2708909 left	43.26%
Xanthomonas axonopodis (all above are Proteobacteria, class Gammaproteobacteria)	3161 right	0.06%	1876940 right	36.27%
Azospirillum sp. (Proteobacteria (α), order Rhodospirillales)	179561 left	5.42%	1595259 right	48.17%
Caulobacter crescentus (Proteobacteria (α), order Caulobacterales)	166112 right	4.14%	1744542 right	43.43%
Azoarcus sp. (Proteobacteria (β), order Rhodocyclales)	32299 right	0.61%	1182414 left	22.49%
Burkholderia cepacia (Proteobacteria (β), Burkholderiales)	160154 right	4.62%	859834 left	24.82%
Campylobacter jejuni (Proteobactria, class Epsilonproteobacteria)	635 right	0.04%	653170 left	40.12%
Group 2				

Myxococcus xanthus (Proteobacteria (Δ), order Myxococcales)	310304 right	3.40%	872133 right	9.54%
Bacteroides thetaiotaomicron (Bacteroidetes, family *Bacteroidaceae*)	246107 right	3.93%	2199265 left	35.10%
Bacteroides fragilis (Bacteroidetes, family *Bacteroidaceae*)	155636 right	2.99%	2420574 left	46.50%
Rickettsia prowazekii (Proteobacteria (α), order Rickettsiales)	382412 left	34.40%	250129 right	22.50%
Neisseria gonorrhoeae (Proteobacteria (β), order Neisseriales)	412632 left	19.16%	616174 right	28.61%
Streptococcus pneumoniae (Firmicutes, family *Streptococcaceae*)	786742 left	37.25%	951471 right	45.05%
Streptococcus pyogenes (Firmicutes, family *Streptococcaceae*)	521569 left	27.53%	897992 right	47.40%
Flavobacterium columnare (Bacteroidetes, class Flavobacteria)	448681 left	14.19%	1240595 right	39.22%
Prevotella intermedia (Bacteroidetes, family *Prevotellaceae)*	267997 left	12.64%	697114 right	32.89%
Sphingobacterium sp. (Bacteroidetes, class Sphingobacteriia)	2928149 right	47.03%	2182864 left	35.06%
Porphyromonas gingivalis (Bacteroidetes, family *Porphyromonadaceae)*	552805 left	23.59%	873251 left	37.26%
Deinococcus radiodurans (Deinococcus-Thermus, class Deinococi)	911819 right	34.45%	714641 left	26.88%
Acidobacterium capsulatum (Acidobacteria, class Acidobacteria)	685949 right	16.62%	33951 left	0.82%
Group 3				
Geobacter sulfurreducens (Proteobacteria (Δ), order Desulfuromonadales)	1831 right	0.05%	downstream	
Pelobacter carbinolicus (Proteobacteria (Δ), order Desulfuromonadales)	2057 right	0.06%	downstream	

Streptomyces coelicolor (Actinobacteria, class Actinobacteria)	3994 left	0.05%	downstream
Mycobacterium tuberculosis (Actinobacteria, class Actinobacteria)	2643 right	0.06%	downstream
Micrococcus luteus (Actinobacteria, family *Micrococcaceae*)	2843 right	0.11%	downstream
Clostridium tetani (Firmicutes, class Clostridia)	2472 left	0.09%	downstream
Lactobacillus brevis (Firmicutes, family *Lactobacillaceae*)	2557 right	0.11%	downstream
Enterococcus faecalis (Firmicutes, family *Enterococcaceae*)	5883 right	0.20%	downstream
Listeria monocytogenes (Firmicutes, order Bacillales)	3776 right	0.13%	downstream
Bacillus subtilis (Firmicutes, order Bacillales)	2546 right	0.06%	downstream
Spirochaeta thermophila (Spirochaetes, class Spirochaetia)	941 left	0.38%	downstream
Fusobacterium nucleatum (Fusobacteria, class Fusobacteriia)	3170 left	0.15%	downstream
Borrelia burgdorferi (Spirochaetes, class Spirochaetia)	1268 left	0.14%	downstream
Mycoplasma haemofelis (Tenericutes, class Mollicutes)	33798 right	2.94%	downstream
Group 4			
Chlamydia psittaci (Chlamydiae, class Chlamydiia)	573957 left	48.97%	downstream
Chlamydia trachomatis (Chlamydiae, class Chlamydiia)	504740 left	48.42%	downstream
Waddlia chondrophila (Chlamydiae, class Chlamydiia)	1044601 right	49.36%	downstream

Table 5.1 Relative positions of *gyrB* and *gyrA* across bacterial species.

5.3 Discussion

In an effort to learn more about the principles underlying the structuring of bacterial chromosomes and gene regulation, the conserved positions of some genes along chromosomes of different bacteria became obvious. Many recent studies (Sobetzko *et al.*, 2012; Slager & Veening, 2016; Sobetzko, 2016; Yubero & Poyatos, 2020; Soler-Bistué *et al.*, 2017) recognised this and started to cast light on the significance of chromosomal gene positions, finding interesting conserved patterns linked to the organisation of gene order. At the same time, some other studies (Brambilla & Sclavi, 2015; Gerganova *et al.*, 2015; Fitzgerald *et al.*, 2015) have questioned the importance of gene order by focusing on nucleoid-associated proteins (NAPs) – global gene regulators and architectural determinants of nucleoid, by altering the chromosomal positions of the promoters, ORFs and/or the entire genes that encode the NAP of interest. The physiological significance of such genetic manipulations, the expression of the repositioned genes in a new chromosomal environment and effects on the regulon of the affected NAP were the major characteristics that were considered. This study is an attempt to contribute to this knowledge by altering chromosomal positions of the major DNA topology modulator, DNA gyrase.

DNA gyrase is composed of two subunits – GyrA and GyrB, that are assembled into the functional enzyme in an A_2B_2 configuration. The subunits are encoded by the separately positioned genes *gyrA* and *gyrB* (Fig. 1.1). This arrangement is present not only in *Salmonella*, but is conserved among gammaproteobacteria (Sobetzko *et al.*, 2012) and, in fact, is present among more clades (Table 5.1). In many other bacteria (Table 5.1), these genes are arranged in the *gyrBA* operon. To study the importance of the geographically dispersed arrangement of the gyrase genes, a novel kind of chromosomal repositioning was carried out – construction of an artificial operon. This arrangement was chosen in an attempt to justify the dispersed locations of *gyrA* and *gyrB*, as expressing subunits of the same protein from the same operon seems more effective. However, the current positions of *gyrA* and *gyrB* in *Salmonella* and other bacteria were established during evolution, meaning that a selective pressure exists for the present arrangement. In both *Salmonella* and *E. coli gyrA* and *gyrB* are similarly regulated – activated by DNA relaxation

(Menzel & Gellert, 1987) and repressed by FIS (Keane & Dorman, 2003; Schneider *et al.*, 1999). This implies that under the same stimuli both subunits are expressed at the similar level, and it was shown at the gene expression level (Fig. 5.7). It might be an advantage, as similar expression of the GyrA and GyrB subunits would allow to assemble the functional enzyme. However, the selective pressure that favoured against the co-expression of *gyrA* and *gyrB* genes from the same promoter in bacteria such as *Salmonella* and *E. coli* is not obvious.

To create the *gyrBA* strain, the ORF of *gyrA*, including a 20 b.p. region upstream that contained a ribosome binding site, was inserted downstream of *gyrB*. The original *gyrA* gene was deleted to avoid having two copies of *gyrA* in the same organism. No *gyrAB* operon is found in bacteria, however, its construction was also attempted. The ORF of *gyrB*, including a 20 b.p. region upstream that contained a ribosome binding site, was successfully inserted downstream of *gyrA*. Attempts to delete the *gyrB* copy from its native locus failed. The resulting *gyrAB_gyrB* strain with two *gyrB* copies was not used in further experimentation since it would be difficult to attribute any of the data that were obtained to the effect of the *gyrAB* operon or to the effect of the extra *gyrB* copy.

The strong impediment to deleting *gyrB* from its original locus suggests the particular importance of the position of this gene. The evidence for the differences between the two genes encoding DNA gyrase was searched that may hint at the reason for the importance of the position of *gyrB* but not the position of *gyrA*. The *gyrB* and *gyrA* genes were found to respond differently to treatment with DNA gyrase inhibitors (Neumann & Quiñones, 1997). In particular, the expression of both *gyrA* and *gyrB* can be induced by coumarins (inhibitors of the GyrA subunit), while only *gyrA* is induced by quinolones (inhibitors of the GyrB subunit). This is despite DNA relaxation occurring in response to both classes of gyrase-inhibiting drugs. This suggests that *gyrB* has different sensitivity to relaxation than *gyrA*. Another characteristic that is different between these genes is the degree of sequence conservation. FIS regulation is important for *gyrA* and *gyrB* in both *E. coli* and *Salmonella.* FIS regulatory binding sites are located upstream from the ORFs of both genes

(Schneider *et al.*, 1999; Keane & Dorman, 2003). However, the 5' region of *gyrA* is significantly diverged between these bacteria, while the 5' region of *gyrB* is, in contrast, highly conserved (Keane & Dorman, 2003). Various mutual arrangements of *gyrA* and *gyrB* were found (Table 5.1), but the relative position of *gyrB* seems to be more conserved than that of *gyrA*. The immediate genetic environment of both genes in bacteria listed in Table 5.1 was studied. One distinct pattern was found – homologues of *dnaA* (encoding chromosomal replication initiation protein DnaA), *dnaN* (encoding the beta subunit of DNA polymerase III) and *recF* (encoding the DNA repair protein RecF) or at least one of these three genes are found directly upstream from *gyrB* gene in all bacteria, where *gyrB* is encoded in the immediate vicinity of the origin, such as most bacteria of groups 1 and 3 (Table 5.1). This is true of most bacteria where *gyrB* is in the immediate vicinity of the origin and *gyrA* is located either further away than about 20% of a chromosome or is a part of the *gyrBA* operon. Bacteria with the *gyrBA* operon not in the immediate vicinity of the origin, such as *Chlamydia psittaci*, and bacteria with separate positioning of the two genes but *gyrA* closer than about 20%, such as *Myxococcus xanthus*, and some Bacteroidetes that satisfy the gene positional parameters characteristic of the group 1 possess *gyrB* with its genetic environment non-conserved. Bacteria of groups 2 and 4 do not have a conserved genetic environment around *gyrB*. The conservation of the genetic environment 5' from the *gyrB* is more important than the subjective proximity to the origin, therefore, it was used as a parameter to decide grouping in the table 5.1. The frequent association of *gyrB* with *dnaA*, *dnaN* and *recF* genes; higher conservation of *gyrB* position in comparison to *gyrA*; the expressional behaviour of *gyrB* in response to quinolones and the inability to delete *gyrB* from the *gyrAB_gyrB Salmonella* strain – all of this evidence indicates that the position of *gyrB* but not *gyrA* is essential in many bacteria.

Following its construction, the whole genome of the *gyrBA* strain was sequenced and was found to contain two SNPs. One of them was T59A in *dxr*, a gene that encodes 1-deoxy-D-xylulose 5-phosphate reductoisomerase and seemed to be not important in the context of DNA gyrase biology. This is because (1) the SNP occurred in the region of DXR that carries no enzymatic

activity – and (2) *dxr* was not found to be involved in any known interactions with DNA gyrase according to the STRING database (Szklarczyk *et al.*, 2015). The second SNP seemed to be more critical, as it occurred directly in the artificial *gyrBA* operon. Although the R6G SNP in the repositioned *gyrA* ORF was in a region of the GyrA subunit that is devoid of any function, it was not desirable to have any SNPs in the chromosomal region that was in the focus of the study. Therefore, the *gyrBA* strain was constructed again using the same methodology and the new construct was found to be free of SNPs after whole genome sequencing. In this work the *gyrBA* strain with the two SNPs is referred to as the *gyrBA**. It should be noted that the SNP in *gyrA*, but not the SNP in *dxr* was the reason for reconstruction of the *gyrBA* strain.

The ability of DNA gyrase to introduce negative DNA supercoils and to relax positive supercoils allows this enzyme to affect multiple cellular processes indirectly. DNA gyrase's major function in the cell cycle is relieving torsional stress caused by positive supercoil accumulation in front of the moving replisome (Liu & Wang, 1987; Chong *et al.*, 2014). This implies that alterations in DNA gyrase processivity may affect the cell cycle and thus morphological and growth parameters. Growth, morphology, and competitive fitness of the *gyrBA* strain were all found to be identical to those of the WT (Fig. 5.1, Fig. 5.3, Fig. 5.4, Fig. 5.5). This indicates that DNA gyrase, produced from the operon, can perform sufficiently well so as to cause no adverse effect on the cell cycle in *Salmonella* and that the *gyrBA* operon does not constitute a burden for the cell. When the growth of the *gyrBA** strain was tested, it was discovered to be slower, as compared to the WT, and the competitive fitness of *gyrBA** was also lower (Fig. 5.2, Fig. 5.5). Since the experiments with the *gyrBA** strain were performed before it was sequenced (and the presence of SNPs discovered), the *gyrBA* operon was thought to be responsible for the changes. However, this hypothesis was refuted after reconstructing the *gyrBA* strain and remeasuring its growth characteristics and competitive fitness. Moreover, the SNP-containing *gyrBA* operon from the *gyrBA** strain was transduced into a clean genetic background to distinguish between the effects of the two SNPs (in this way the SNP in *dxr* was removed and the SNP in *gyrA* was retained) and its growth was remeasured. The transductant with the *gyrBA** operon, (with its

gyrA SNP), in a clean genetic background was found to grow identically to the WT (Fig. 5.2c). Thus, the SNP in *dxr* was causing growth and competitive fitness deficiency, although it had been deemed unimportant at first. Similarly, motility was measured in the *gyrBA** strain and was found to be decreased in comparison to the WT, but in the *gyrBA* strain the difference was eliminated (Fig 5.6). The effect of the SNP in *dxr* on the phenotypes that have been investigated could be explained by the essentiality of *dxr* (gene that encodes DXR – an enzyme that is utilized in the committed step of the MEP pathway that synthesizes essential isoprenoid compounds in *Salmonella*) (Takahashi *et al.*, 1998). No mutations that inactivate *dxr* are tolerated and all the mutants of the genes downstream of *dxr* in the MEP pathway displayed growth deficiencies (Cornish *et al.*, 2006). These findings highlight the importance of avoiding any non-synonymous SNPs during strain construction, as the effect of the SNP can be mistaken for the effect of the desired mutation.

Repositioning of the *gyrA* ORF to the new chromosomal location interrupted the existing transcription regulatory networks. In previous studies where ORFs of different nucleoid-associated proteins were exchanged, gene expression patterns of the resulting hybrid genes were rescheduled or exchanged (Fitzgerald *et al.*, 2015). Although, the transcriptional regulation of *gyrA* and *gyrB* genes in *Salmonella* is highly similar (Keane & Dorman, 2003), there are differences between these promoters. First difference is the lack of conservation of the 5' region of *gyrA* as opposed to the 5' region of *gyrB* among bacteria, mentioned above (Keane & Dorman, 2003). Second significant difference between the P_{gyrA} and P_{gyrB} is that the gene expression of *gyrB* but not that of *gyrA* can be mediated by RpoS sigma factor of RNA polymerase in both *E. coli* and *Salmonella* (Lévi-Meyrueis *et al.*, 2014; Maciag *et al.*, 2011). This may be one of the reasons that resulted in the inability to delete *gyrB* from the *gyrAB_gyrB* strain. Nevertheless, no major change of gene expression patterns of both *gyrA* and *gyrB* genes was detected in the *gyrBA* strain (Fig. 5.7). This could be because even though only *gyrB* can be transcribed by RpoS, both *gyrA* and *gyrB* can be transcribed by RpoD sigma factor, meaning that the *gyrA* ORF, under the regulation of the P_{gyrB} in the *gyrBA* operon, retained native-like regulation at the new chromosomal position. Therefore, in

the *gyrBA* strain the *gyrA* ORF is likely to receive highly similar regulatory signals to those generated by its own promoter. In the *gyrAB_gyrB* strain the *gyrB* ORF is placed under the control of the P_{gyrA} and is therefore cannot be transcribed by RpoS as in the WT. This was likely one of the reasons for the inability to delete *gyrB* from its original locus.

If the *gyrBA* operon configuration of *gyrA* and *gyrB* genes affects properties of the DNA gyrase holoenzyme, it might manifest through altered interactions of DNA gyrase with other molecules. An example of such molecules is antibiotics that target DNA gyrase. There are two different classes of drugs that differ in their mode of action on DNA gyrase: those that act via the GyrA subunit (quinolones), and via the GyrB subunit (aminocoumarins). The ability of these antibiotics to inhibit growth of the WT and the *gyrBA* strains was tested. MIC_{90} concentrations of the quinolones were identical in both the *gyrBA* strain and the WT (Fig. 5.8c and Fig. 5.8d). While the resistance of the *gyrBA* strain to aminocoumarins was altered compared to the WT (Fig. 5.8a and Fig. 5.8b). This suggests that the creation of the *gyrBA* operon resulted in production of DNA gyrase that has altered molecular interactions with antibiotics of the aminocoumarin class. Alternatively, uptake of these antibiotics is altered in the *gyrBA* strain. Since the genome of the *gyrBA* strain is identical to that of the WT, except the presence of the *gyrBA* operon (Section 5.2.2), no changes to the structure of the GyrA and GyrB subunits were introduced in the *gyrBA* strain. However, the creation of the operon might have affected the efficiency of assembly of DNA gyrase, this in turn may have influenced the molecular interactions of the aminocoumarins with the GyrB subunit. The opposing response of the *gyrBA* strain to novobiocin and coumermycin can also be explained by the altered gyrase assembly, since these drugs differ in their binding properties – novobiocin binds the GyrB subunit and blocks one ATP-binding site, while coumermycin binds two GyrB subunits blocking two ATP-binding sites (Vanden Broeck *et al.*, 2019b; Bush *et al.*, 2015).

To measure processivity of DNA gyrase when it is expressed from the *gyrBA* operon in *Salmonella in vivo*, global DNA supercoiling of the *gyrBA* strain was determined, using the pUC18 reporter plasmid, and comparing it to that of the WT. DNA in the *gyrBA* strain was more relaxed than that in the WT when grown

in LB (Fig. 5.11). However, in the N-minimal medium, in contrast, the *gyrBA* strain had its genome more supercoiled than the WT (Fig. 5.12). Since the extent of the global DNA supercoiling of the *gyrBA* strain depended on the growth medium, it cannot be stated that the *gyrBA* strain supercoiled its genome more or less efficiently than the WT. Instead, the *gyrBA* had different supercoiling set points that had to be adapted to the growth conditions. It would be interesting to look at the local DNA supercoiling using supercoiling sensors (Rovinskiy *et al.*, 2019) or treatment with trimethyl psoralene (Lal *et al.*, 2016) as the global supercoiling measurements capture average supercoiling density of the genome only. The smaller difference in global DNA supercoiling between the WT and the *gyrBA* in the minimal medium with low Mg^{2+} concentration, compared to the high Mg^{2+} (Fig. 5.12), suggests that DNA gyrase that is expressed from the *gyrBA* operon may have altered interactions with Mg^{2+} ions at the metal-binding topoisomerase-primase (TOPRIM) domain of *gyrB* (Fig. 1.5).

The DNA supercoiling differences found in the *gyrBA* strain may affect global gene expression, including that of genes in the *Salmonella* Pathogenicity Islands (SPIs) – the genetic regions that express type III secretion systems and effector proteins that enable entry into epithelial cells and survival inside the vacuole of a macrophage. Although both SPI-1 and SPI-2 gene expression are regulated by DNA relaxation (Cameron & Dorman, 2012) and that DNA was more relaxed in the *gyrBA* strain during growth in LB (Fig. 5.11), the expression of both SPIs in the *gyrBA* was found to be indistinguishable from that in the WT when grown in LB broth (Fig. 5.9). However, the significant decrease of SPI-2 expression was discovered with the aid of growth in a minimal medium with the low concentration of Mg^{2+} ions (Fig 5.10d). These conditions were designed to mimic the environment that *Salmonella* encounters inside the *Salmonella*-containing vacuole (SCV) of a macrophage and the decrease in SPI-2 expression indicates at potential deficiency in the ability of the *gyrBA* strain to survive within a macrophage. To study the performance of the *gyrBA* strain inside its major niche during infection cycle, infection assays of RAW 264.7 macrophages were performed. To further vary SPI-1 and SPI-2 expression dynamics during infection, different modes of *Salmonella* entry into macrophage

were utilized: SPI-1-promoted active invasion and passive entry via phagocytosis followed by SPI-2-promoted intracellular survival. In line with the decrease in the SPI-2 expression observed in the conditions that mimic SCV, there was a decrease in the ability of the *gyrBA* strain to survive and replicate within a murine macrophage at 16 h compared to the WT (Fig. 5.14b). There was also an increase in the rate of entry of the *gyrBA* strain into a murine macrophage (Fig. 5.14a). The reason for this increase is not immediately obvious, since there were no changes in SPI-1 expression observed. However, it can be noticed that the growth medium used in this experiment did not allow for the detectable expression of SPI-1 (Fig. 5.10), and the growth in LB broth did not reveal the SPI-1 expression alteration in the *gyrBA* strain (Fig. 5.9). Thus, the possibility of detecting the SPI-1 expression differences between the WT and the *gyrBA* strains exists at other conditions that were not tested. At the same time, there were no differences in entry into murine macrophages via phagocytosis and the subsequent survival between the *gyrBA* and the WT strains. The presence of the replication defect in the *gyrBA* strain that followed the SPI-1 mediated entry and its absence when bacteria were phagocytosed confirms that the SPI expression dynamics are distinct between the two models of entry.

Experiments carried out in this work hinted at the importance of the position of *gyrB* gene and revealed some alterations in the physiology of the *gyrBA* strain. However, the main question of the significance of the separate relative positioning of *gyrA* and *gyrB* in some bacteria, as opposed to the *gyrBA* operon in the other bacteria remained unanswered. Thus, further analysis of the mutual positions of the genes that code for DNA gyrase was performed.

Global DNA supercoiling is not uniform. At the stationary growth phase circular chromosome has a high level of negative supercoiling near the terminus that gradually decreases toward the origin (Lal *et al.*, 2016), while at the mid-exponential phase of growth supercoiling is equal throughout the chromosome. At the same time, it was shown that at the exponential phase of growth the DNA gyrase binding sites are enriched near the origin of chromosomal replication (Jeong *et al.*, 2004; Sutormin *et al.*, 2019). These findings may seem

contradicting, however, when the exponential and stationary phase supercoiling levels were directly compared, negative supercoiling was shown to be increasing more at the *oriC* at the exponential phase (Lal *et al.*, 2016). This is because bacterial chromosome is more supercoiled at the exponential than at the stationary growth phase (Balke & Gralla, 1987), thus to neutralize the DNA supercoiling gradient at the exponential phase, the DNA gyrase activity at the *oriC* is required. This agrees with the global distribution of gyrase at the exponential phase (Sutormin *et al.*, 2019; Jeong *et al.*, 2004). Recent mapping of gyrase cleavage sites revealed that these are concentrated not only near the origin of chromosomal replication but also downstream of highly transcribed operons (Sutormin *et al.*, 2019). This confirms that DNA gyrase acts upon positively supercoiled DNA that is located near the origin and resolves positive supercoiling associated with transcription (Chong *et al.*, 2014). The transcribing RNA polymerase not only induces local DNA supercoiling, but also provides chromosome domain boundaries that constrain supercoils (Booker *et al.*, 2010; Higgins, 2014). However, global distribution of supercoiling-sensitive genes on a chromosome is not limited to the origin and terminus-proximal regions, but instead is randomised in Enterobacteriaceae (Sobetzko, 2016), or organised in clusters in *Streptococcus pneumoniae* (Ferrándiz *et al.*, 2010). Such a diverse distribution does not allow the global DNA supercoiling to regulate the supercoiling-sensitive genes, instead the required supercoiling stimulus is provided by the local supercoiling generated by transcribing RNA polymerase (Liu & Wang, 1987; Le *et al.*, 2013). This means that transcriptional activity and orientation of neighbouring genes is important to determine the level of local supercoiling that acts on a supercoiling-sensitive promoter. The local supercoiling, although is acted upon by topoisomerases to balance it, can be diffused freely affecting transcription of supercoiling-sensitive promoters within approximately a 10 kb range. The effect of local DNA supercoiling created by highly transcribing genes has most significant influence on immediate gene neighbours (Sobetzko, 2016). 10 kb range corresponds to previously determined chromosomal microdomain size (Postow *et al.*, 2004).

The promoters of *gyrA* and *gyrB* are supercoiling-sensitive, resulting in DNA gyrase upregulation by the relaxed state of DNA (Menzel & Gellert, 1987). The

local genetic environment of *gyrA* in bacteria where it is positioned separately from *gyrB* (Table 5.1) is not conserved. It is also hard to see from where the P_{gyrA} sources DNA relaxation required for its expression. Transcription from the neighbouring convergent genes would provide the relaxation required (Sobetzko, 2016), however, the separately-positioned *gyrA* genes, in the majority of bacterial species tested, are located between divergent operons that instead supply *gyrA* promoter with negative supercoiling during transcription. Therefore, no logical patterns in chromosomal position of the separately positioned *gyrA* among different bacterial species were found. Similarly, no patterns in the chromosomal position of *gyrB* when it is away from *oriC* and *dnaA-dnaN-recF* operon(s) were found.

Genes *dnaA, dnaN*, *recF* can be organised in one or more transcription units in different bacteria, always arranged in a tandem (co-oriented) with *gyrB*, transcribed at high level but separately from *gyrB* (Macián *et al.*, 1994). This means that the 5' *dnaA-dnaN-recF* combination of the co-oriented genes supplies P_{gyrB} with relaxed local supercoiling input (Sobetzko, 2016) that is required for the upregulation of *gyrB* expression. This creates high selective pressure for *gyrB* or *gyrBA* to remain at this chromosomal position.

To summarize, the detrimental effect associated with the *gyrBA* operon in *Salmonella*, was associated with the different ability of the bacterium that harbours it to supercoil its DNA. This led to the altered expression of SPI-2 and subsequent decrease in the ability of the *gyrBA* strain to survive and replicate inside a mammalian macrophage. Other tested characteristics, compared to the WT, remained unchanged. This showed that the *gyrBA* operon was tolerated surprisingly well by *Salmonella* that naturally carries *gyrA* and *gyrB* genes in the dispersed chromosomal locations. However, the altered ability of the *gyrBA* strain to supercoil its chromosome suggests that the global gene expression patterns of the *gyrBA* strain at the mRNA production and, subsequently, at the protein synthesis level differ, due to the gene regulatory role of DNA supercoiling. Regarding the gene positioning, multiple arguments were found that explain the conserved *oriC*-proximal location of *gyrB* or the *gyrBA* operon among different bacteria, but the justification for separate positioning of *gyrA* in other bacteria remains obscure.

Chapter 6:

General discussion

6.1 General discussion.

6.1.1 Background

The bacterial genome may consist of one or more circular or linear chromosomes and plasmids organised in the structure termed the nucleoid. This mass of DNA is not amorphous but is instead organised at multiple levels with the aid of DNA binding proteins, RNA and DNA supercoiling. The resulting structure is highly dynamic, allowing the genome to replicate efficiently and to facilitate other DNA transactions that ensure the optimal performance of the bacterium in its environment (Dorman, 2013). To ensure the correct and timely segregation of daughter chromosomes, different regions of the chromosome must interact with each other in a coordinated manner. This is organised at the level of macrodomains – chromosomal regions that do not interact with each other (Valens *et al*., 2004; Dame *et al*., 2020). Not only the order of macrodomains, but the order of individual genes on the chromosome is important. This can be seen from the high degree of conservation of gene positions between related bacteria (Sobetzko *et al*., 2012; Yubero & Poyatos, 2020). Nevertheless, the exact reasons for the existing chromosomal gene positions are often obscure. One strategy to obtain an insight into the significance of the chromosomal position of a gene is alter it and attempt to interpret the associated phenotypes.

The two studies presented in this thesis were inspired by the same philosophy and shared similar methodology. They primarily aimed at identifying the significance of the chromosomal positions of genes that encode the major gene regulatory proteins IHF and DNA gyrase. With regard to this primary aim, both studies yielded a similar result – the original gene positions are not essential for survival and their alteration does not result in substantial deleterious effects in *Salmonella*, although there were changes found with mild but diverse outcomes. However, unanticipated results and observations of these studies led to curious and largely unexpected conclusions in each of the two studies. I will first discuss results that directly relate to the significance of the chromosomal gene positions.

6.1.2 Physiological aspects of *Salmonella* not influenced by the chromosomal repositionings of genes

Following construction of the strains with the translocations of the genes that encode IHF and DNA gyrase, whole genome sequencing of these strains was performed to ensure that any phenotypes observed were due to the repositionings and not due to additional unwanted mutations (Sections 3.2.2 and 5.2.2). No such mutations were detected. It soon became evident that none of these strains differ from the WT SL1344 in their growth characteristics (Figs. 3.2-3.3 and Figs. 5.1-5.3) and morphology (Figs. 3.4-3.5 and Fig. 5.4). This highlights the fact that the cell cycle of *Salmonella* is not affected by the repositionings, while complete deletion of any of the *ihf* genes affects both growth dynamics and morphology (Figs. 3.2-3.5). Deletion of the *gyr* genes is not possible since DNA gyrase in essential for *Salmonella*, thus, if the *gyrBA* operon was to affect DNA gyrase function in a serious enough way to affect the cell cycle, the bacterium would probably be not viable.

Therefore, repositioning of genes such as *ihfA*, *ihfB* of *gyrA* has much milder effect than deletion of these genes. This finding is also consistent with the observation that there is little or no difference in expression of *Salmonella* pathogenicity islands SPI-1 and SPI-2 between the repositioned strains and the WT, while the Δ*ihfA* and Δ*ihfB* strains displayed a substantial decrease in expression of both SPIs (Fig. 3.8 and Fig. 5.9). Competitive growth of two co-cultured bacterial strains in the same flask can detect fine differences in fitness between the strains. In prior work with strains with chromosomal gene repositionings a decrease or even an increase in fitness can be observed upon varying growth conditions (Fitzgerald *et al.*, 2015; Gerganova *et al*., 2015). In the present study, there were no differences in competitive fitness between the WT and any of the repositioned strains at the standard growth conditions (Fig. 3.6a and Fig. 5.5). However, an increase of osmotic pressure in the growth medium caused the *ihfAB* strain to be outcompeted by the WT (Fig. 3.6b). This highlights the fact that the repositioned strains are not all physiologically equivalent, and a difference can be detected even between the strains with different configurations of the artificial IHF operon. This is further supported by a small difference found between the IHF operon strains, with SPI-1 expression

being lower in the *ihfBA* strain but not in the *ihfAB* strain (Fig. 3.8). Change of growth medium during SPI expression measurements to the one resembling conditions inside a SCV revealed a SPI-2 expression defect in the *gyrBA* strain that was not detectable at the standard growth conditions (Fig. 5.10).

The virulence of the gene repositioned strains was tested using different modes of *Salmonella* entry into a murine macrophage (phagocytosis or SPI-1-induced) (Drecktrah *et al.*, 2006). Variable results were obtained, but here I will describe those that did not alter between the strains. No difference in long-term bacterial survival and replication between the OrfSwap$^{ihfA\text{-}ihfB}$ and the WT strains was found when the phagocytosis-mediated entry mode was used (Fig. 3.9b). This finding is in line with the lack of change in SPI-2 expression in the OrfSwap$^{ihfA\text{-}ihfB}$ strain (Fig. 3.8). There were also no differences in entry and the subsequent survival between the *gyrBA* and the WT strains using the phagocytic entry model (Fig. 5.13). Similarly, no difference in the SPI-1-mediated entry was detected in the OrfSwap$^{ihfA\text{-}ihfB}$, *ihfAB* or *ihfBA* strains (Fig. 3.10), in line with no detectable differences in SPI-1 expression in those strains (Fig. 3.8).

Expressing *gyrA* and *gyrB* from the operon did not affect transcription of either gene (Fig. 5.7). This is despite a large change in gene dosage of *gyrA* due to relocation from a Ter- to an *oriC*-proximal position (Fig. 2.1). Since DNA gyrase is essential for *Salmonella*, it is important to produce the subunits of this enzyme in a precise 2:2 ratio for correct assembly. It is possible that the effects of even such a large change in gene dosage, as the one arising due to the *gyrBA* operon construction, become neutralized by tight post-transcriptional control. I conclude that *Salmonella* can tolerate *ihfA*, *ihfB* and *gyrA* chromosomal gene repositionings very well with only negligible effects on growth and SPI expression. However, a more refined ability to manipulate competitive fitness and SPI expression by small changes in growth conditions makes it likely that future testing in other infection-relevant conditions will reveal further details.

6.1.3 Physiological aspects of *Salmonella* that are affected by altered gene position

Many phenotypic differences were discovered between the WT and the strains with the chromosomal repositionings. IHF is a major regulator of motility in *Salmonella* (Mangan *et al.*, 2006) and the OrfSwap$^{ihfA\text{-}ihfB}$ strain displayed a motility decrease, while the *ihfAB* and *ihfBA* strains preserved the WT-like motility (Fig. 3.7). Although DNA gyrase has not been implicated in flagellar motility, a small motility decrease was observed in the *gyrBA* strain that is not likely to be biologically significant (Fig. 5.6).

The OrfSwap$^{ihfA\text{-}ihfB}$ strain was found to be phagocytosed by macrophage more efficiently than the WT (Fig. 3.9a). This could be due to alterations in the bacterial cell surface composition in the OrfSwap$^{ihfA\text{-}ihfB}$ strain (Murray *et al*., 2006). The proteomic study identified numerous proteins that were differentially expressed in this strain and the WT and these included proteins involved in transport and outer membrane composition (Table 4.2). However, the proteomic study was performed at growth conditions that were dissimilar with the conditions that precede phagocytosis and was aimed to sample the proteome of *Salmonella* at the stage of growth when IHF is expressed at its highest level (late exponential-early stationary transition growth stage in LB, see below). Therefore, the enhanced phagocytosis cannot be explained with the results of the global protein production study. There was a decrease in the ability of the *gyrBA* strain to survive and replicate within a murine macrophage (Fig. 5.14b) that can be explained by the decrease in the SPI-2 expression observed in the conditions that mimic the SCV environment (Fig. 5.10). There was also an increase in the rate of entry of the *gyrBA* strain into a murine macrophage (Fig. 5.14a). This increase was unexpected, since in the *gyrBA* strain no changes in SPI-1 expression were observed (Fig. 5.9 and Fig. 5.10). However, the conditions that were used to mimic the SCV were not favourable for any SPI-1 expression (Fig. 5.10), thus, it is possible that a SPI-1 expression increase in the *gyrBA* strain will be manifested in other growth conditions. Since *Salmonella* invades epithelial cells in the lumen of the small intestine prior to encountering macrophages (Fig. 1.7), it would be interesting to assess the

dynamics of entry of all the repositioned strains into human epithelial cell lines, such as HeLa or Caco-2.

The exchange of *ihfA* and *ihfB* resulted in a novel rescheduled transcription pattern for these genes that cannot be characterised simply as a decrease or an increase (Fig. 4.2). The creation of the *ihfAB* and *ihfBA* operons resulted in transcription patterns that were lower in magnitude than those of the WT, but otherwise closely resembled the WT (Fig. 4.3). This is further evidence (in addition to the effects on motility described above) indicating that the operonic arrangement of *ihfA* and *ihfB* has milder effects on *Salmonella* than the reciprocal exchange of these genes.

Antibiotics of aminocoumarin and quinolone classes bring about their antimicrobial activities by targeting DNA gyrase (Lewis *et al*., 1996; Drlica & Zhao, 1997). It was important to determine if the expression of DNA gyrase from an operon resulted in an altered response to these drugs. There were no differences in MICs of quinolones between the WT and the *gyrBA* strains. Interestingly, the response to different aminocoumarins at subinhibitory concentrations of antibiotics was opposing: *gyrBA* was less resistant to coumermycin, but more resistant to novobiocin (Fig. 5.8). The changes in MICs indicate that molecular interactions of aminocoumarins with DNA gyrase in the *gyrBA* strain are altered. Since coumermycin and novobiocin act via binding to GyrB but in different ways (Vanden Broeck *et al*., 2019b; Bush et al., 2015), the opposite response to these antibiotics may mean that assembly of DNA gyrase in the *gyrBA* strain is altered. A positive effect of the operon on assembly is intuitively appealing, since the adjacent location of *gyrA* and *gyrB* in the operon is radically different from the WT locations, where these genes are associated with opposite poles of the chromosome. Alternatively, a change of global gene expression in the *gyrBA* strain brought about a change in the rates with which the two drugs enter the cell.

The primary and unique function of DNA gyrase is the introduction of negative supercoils into DNA (Champoux, 2001). Therefore, global DNA supercoiling was tested in the *gyrBA* strain and compared to that of the WT. It was found to be different between the WT and the *gyrBA* strains, but the direction of this

difference depended on culture conditions. The *gyrBA* strain had its DNA more relaxed than the WT when grown in LB (Fig. 5.11). In the minimal medium, its DNA was, in contrast, more supercoiled (Fig. 5.12). Such opposing patterns of DNA supercoiling do not allow one to conclude unambiguously if the *gyrBA* strain supercoils its DNA to higher or lower extent than the WT. Instead, the *gyrBA* strain appears to adapt to changing growth conditions by adjusting its DNA supercoiling in a way that is different from that of the WT. A further level of complexity comes from the DNA supercoiling response to altering the Mg^{2+} concentration of the medium. While the linking number difference was higher at high Mg^{2+} concentration, it decreased at a low concentration of Mg^{2+} ions – the growth condition that was used to mimic the environment inside the SCV. This suggests that DNA gyrase that is expressed from the *gyrBA* operon may have altered interactions with Mg^{2+} ions at the Mg^{2+} binding TOPRIM domain of GyrB subunit (Fig. 1.5). Alternatively, expression of the magnesium transporters is altered in the *gyrBA* strain. Rather than testing DNA supercoiling in media formulated to resemble the SCV, it would be interesting to measure DNA supercoiling in the *gyrBA* strain that has been directly internalised by a macrophage. While it is known that *Salmonella* relaxes its DNA inside a macrophage (O'Croinin *et al.*, 2006), these results would provide a truly infection-relevant conclusion about the extent to which the *gyrBA* strain is able to supercoil its DNA. Alteration of global DNA supercoiling leads to the consequent alterations in the compaction level of a genome, as well as in global gene expression patterns. Therefore, it would be of a special interest to monitor the picture of global changes that occur in *Salmonella* harbouring the *gyrBA* operon on transcriptional, protein production levels, or both.

A global protein production study was performed to uncover differences between the WT and the OrfSwap$^{ihfA\text{-}ihfB}$ strains. This was done at the exponential-early stationary transition phase of growth (7 h of growth in LB) to maximise IHF expression and, therefore, to gain a better understanding of the significance of chromosomal positions of genes that encode IHF on its global functionality. This proteomics experiment had been planned initially to determine the ratio of IHF subunits in a separate study that questioned the stoichiometry of IHF (see below) but resulted in the elucidation of a list of

differentially expressed proteins as a bonus. There were multiple differentially expressed proteins detected in the OrfSwap$^{ihfA\text{-}ihfB}$ strain that were categorised according to a role and localisation in a cell. More proteins were downregulated (214) than upregulated (84) and this is likely the result of the downregulation of multiple proteins associated with translation (Fig. 4.11). The widespread downregulation of ribosomal proteins means that the overall rates of protein synthesis in the OrfSwap$^{ihfA\text{-}ihfB}$ strain are likely to be decreased. This makes distinguishing between the true effects of the reciprocal ORF exchange of *ihfA* and *ihfB* and the general bias to protein downregulation difficult. On the other hand, since downregulation of the ribosomal proteins is the result of the ORF exchange, all other alterations are similarly caused by it directly or non-directly. Change in the production of some important regulatory proteins was observed and this is likely to produce downstream effects on their regulons (Fig. 4.10). It was expected that many proteins involved in motility, chemotaxis, stress response and virulence would be affected as the genes that express these are regulated by IHF (Mangan et al., 2006; Prieto et al., 2012). In reality, only a few of those were differentially expressed in the OrfSwap$^{ihfA\text{-}ihfB}$ strain (Table 4.2). It should be noted that, in the previous works on IHF, the global transcriptomes were studied using single and double *ihfA*/*ihfB* deletions. The genetic manipulations used in the present work were much more gentle and resulted in the complex reprogramming of IHF's regulation rather than complete inactivation of its subunits. This makes direct comparisons of this work with the earlier studies difficult. The validity of the proteomic experiment was confirmed by Western blotting of the up- and downregulated proteins RpoD and H-NS (Fig. 4.15), as well as, in retrospect, by Western blotting of IhfA and IhfB (Fig. 4.7).

It can be argued that usage of the S_0 parameter, when setting the threshold cut offs that result in smooth hyperboloid thresholds distinguishing statistically significant, differentially regulated peptides, is not ideal (Fig. 4.9). Linear fold-change and p-value thresholds may serve as cut offs that are more robust (Giai Gianetto *et al.*, 2016). However, the widespread usage of S_0 in modern proteomics and its availability with no alternative for the fold-change cut off in

Perseus (the popular proteomics statistical software) (Tyanova *et al.*, 2016) made me choose it for the analysis.

While the OrfSwap$^{ihfA\text{-}ihfB}$ strain had been subjected to intense experimentation, the *ihfAB* and *ihfBA* operon strains were tested only in a few cases and their characterisation remains preliminary, mostly due to time constraints. The artificial operonic arrangement of regulatory genes is interesting due to its novelty – no other studies that dealt with chromosomal gene repositionings approached it by constructing operons (Brambilla & Sclavi 2015; Gerganova *et al*., 2015; Fitzgerald *et al*., 2015; Bogue *et al*., 2020; Mogre & Dorman, in preparation). Findings that *ihfA* and *ihfB* transcription as well as flagellar motility in the strains that harbour artificial operon operons are more WT-like than reciprocal exchange of these genes, make further investigations promising. Additional experimentation with the *ihfAB* and *ihfBA* strains warrants more insights into how the operonic and separate arrangements of *ihfA* and *ihfB* differ.

6.1.4 IHF stoichiometry

The IHF study had an unexpected outcome that is not linked to the position of *ihfA* and *ihfB* on the chromosome. It is related to IHF subunit stoichiometry and was first noticed during the analysis of IhfA and IhfB protein production by Western blotting in the WT and the OrfSwap$^{ihfA\text{-}ihfB}$ strains, engineered with epitope tags to enable immunodetection. The levels of IhfB were found to be higher than those of IhfA and this was best seen at the exponential-to-stationary transition phase of growth when IHF expression peaks (Fig. 4.7c). This difference was further increased in the OrfSwap$^{ihfA\text{-}ihfB}$ strain (Fig. 4.7d). The Western results were not sufficient to claim that there is more IhfB in a cell than IhfA, since epitope-tagged strains were used for this experiment. Although the tags were shown not to have any effect on the transcription of *ihfA*, *ihfB* and IHF regulatory targets (Fig. 4.5 and Fig. 4.6), an effect of the tags on IHF subunit protein production cannot be ruled out. It was, therefore, necessary to determine the IhfA/IhfB ratio in the untagged WT and the OrfSwap$^{ihfA\text{-}ihfB}$ strains. This was achieved by mass spectrometry using the data obtained from the same experiment that assessed the differential expression of proteins between the WT and the OrfSwap$^{ihfA\text{-}ihfB}$ strains. In fact, this experiment was originally conceived to serve the purpose of IhfA/IhfB ratio determination, and the differential expression data were obtained as a bonus. It was discovered, using the iBAQ intensity values to enable comparison of protein intensities within a sample (Schwanhäusser *et al.*, 2011), that there is almost twice as much IhfB in the WT SL1344 as IhfA at 7h (corresponding to the exponential-stationary transition) (Fig. 4.12). These data confirm the Western blotting results for the WT strain. Interestingly, while Western blotting showed that the difference between the IhfA and IhfB was further increased in the OrfSwap$^{ihfA\text{-}ihfB}$ strain, the MS analysis revealed a swich in the IhfA/IhfB ratio in this strain (Fig. 4.12). Since MS is by far the more reliable for this experimental purpose, it can be confidently said that the abundance of IhfB is higher than that of IhfA in the WT strain, while the abundance of IhfA is higher than that of IhfB in the OrfSwap$^{ihfA\text{-}ihfB}$ strain. The disagreement between the Western and the MS results in the OrfSwap$^{ihfA\text{-}ihfB}$ strain is because the protein production in the OrfSwap$^{ihfA\text{-}ihfB}$ strain was affected by tagging (Section 4.2.13). This highlights the fact that the

epitope tagging may sometimes (no effect of tagging on the WT strain was shown in the Section 4.2.13) cause widespread effects on global protein production due to a partial loss of functionality of the tagged protein that cannot be detected at the level of transcription of individual genes. Thus, methods other than epitope tagging should be prioritised in future work; for example, the use of a protein-specific antibody, where possible, or the application of MS analysis. The similar transcription levels of *ihfA* and *ihfB* (Fig. 4.2) and the higher abundance of IhfB (Fig. 4.7) imply that post-transcriptional regulation is responsible for the unequal abundance of IHF subunits. Stability of *ihfA* and *ihfB* mRNA was assessed in the WT and the OrfSwap$^{ihfA\text{-}ihfB}$ strains and revealed that the *ihfB* mRNA molecule is more stable than *ihfA* mRNA in the WT, while the opposite is true in the OrfSwap$^{ihfA\text{-}ihfB}$ strain. This fits well with the results of other experimentation.

The discovery that IhfA and IhfB are not present in a 1:1 ratio was only partly surprising. After establishing that IhfA and IhfB are present in equal amounts in order to assemble into a functional heterodimer (Ditto *et al.*, 1994; Ali Azam *et al.*, 1999) multiple lines of evidence indicate that IHF can function as a homodimer. IhfB, but not IhfA, can efficiently form homodimers *in vitro* (Zulianello *et al.*, 1994). The IhfA and IhfB homodimers have unique, but overlapping, regulons and can form homodimers *in vivo* in IHF single subunit deletion mutants in both *Salmonella* and *E. coli* (Mangan *et al.*, 2006; Prieto *et al.*, 2012). Moreover, MS data showing that IhfB is more abundant than IhfA had long been available and compiled in PaxDB database of protein abundance (Wang *et al.*, 2015) but was not noticed because it was obtained in studies that were not interested in IHF (Adkins *et al.*, 2006; Correia *et al.*, 2014). The study by Prieto *et al*., 2012 provided more evidence that IHF can be a homodimer of IhfB. ChIP-seq of individual IHF subunits revealed that IhfB covers noticeably higher proportion of *E. coli* genome than IhfA. In addition, IhfB binding regions are extensions of IhfA binding regions. Interestingly, ChIP-seq of IhfB in the strain with deleted *ihfA* and the opposite experiment (ChIP-seq of IhfA in the strain with deleted *ihfB*) were not successful since not enough DNA was obtained for the sequencing. This suggests that both IHF subunits are required for successful binding of IHF homodimers to DNA.

The picture shaped with the aid of the results from experimentation relating to IHF stoichiometry complements the hypothesis formulated by Prieto *et al* 2012 and suggests that *ihfA* and *ihfB* are transcribed to produce mRNA in a ratio close to 1:1. The *IhfA* mRNA gets degraded at a faster rate than *ihfB* mRNA and this leads to the synthesis of a relatively higher amount of the IhfB subunit. Then IHF assembles in the form of a heterodimer of IhfA and IhfB, and a homodimer of IhfB, that is formed only in the presence of a nucleating IHF heterodimer to carry out its DNA architectural and gene regulatory roles in the cell (Fig. 6.1). However, formation of a small number of IhfA homodimers cannot be ruled out. In order to complete this picture and to fully account for the different abundances of IHF subunits, the dynamics of IhfA and IhfB protein turnover will have to be monitored to complement the *ihfA* and *ihfB* mRNA degradation experiment. It was previously shown that changes in IHF concentration, KCl concentration and tension applied to DNA can dramatically influence the degree of DNA bending, leading to slight bends in DNA via non-specific binding when IHF is overcrowded (Lin *et al.*, 2012). While a similar IHF-DNA interaction may occur in the case of the cooperative binding of IHF heterodimer and the IhfB homodimer(s), it should be noted that the real mode of DNA bending by such a homomeric complex is not known.

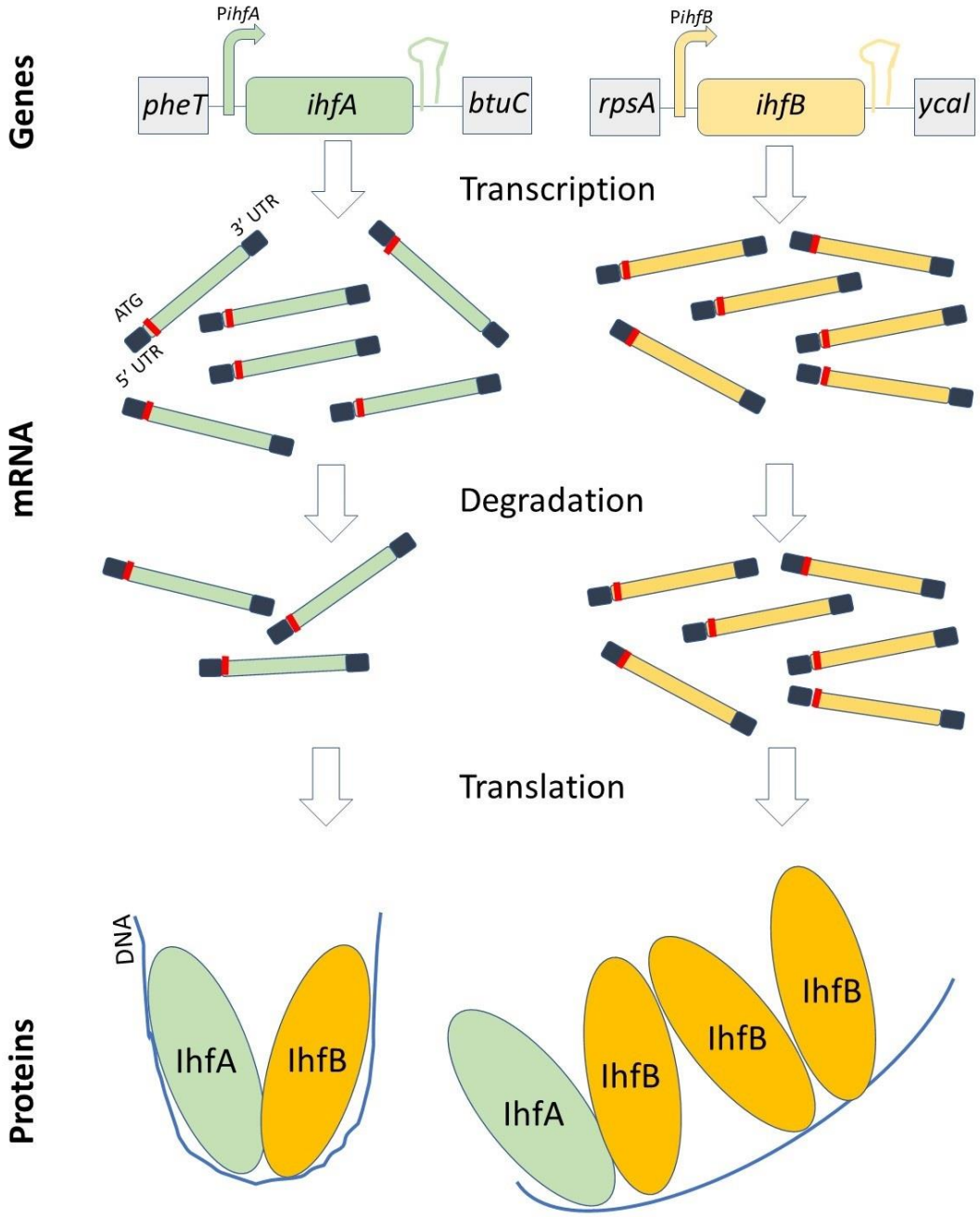


Figure 6.1 IHF, from transcription to DNA binding. In the WT *Salmonella* *IhfA* and *ihfB* are transcribed to produce mRNA in a ratio close to a 1:1. *IhfA* mRNA gets degraded at a faster rate than *ihfB* mRNA and this leads to the synthesis of a higher amount of the IhfB subunit. IHF assembles in the form of a heterodimer of IhfA and IhfB that bend DNA at a sharp angle upon binding. A homodimer of IhfB, is also formed but only in the presence of a nucleating IHF heterodimer. The way in which this complex bends DNA is not known, and a slight bending is only shown for the purpose of illustrating a general principle.

6.1.5 Chromosomal position of *gyrA* and *gyrB* among bacteria

Part of the DNA gyrase study not related to the experimental comparison of the WT and the *gyrBA* arrangements of *gyrA* and *gyrB* genes in *Salmonella*, involved observations of the chromosomal positions of these genes across other bacteria (Section 5.2.13). Representatives of different bacterial phyla were included in the comparison in an attempt to achieve maximum coverage. It was discovered that bacteria can be grouped on the grounds of: 1 – operonic or separate arrangement, 2 – proximity of *gyrB* to *oriC*. As the grouping exercise was carried out, it became evident that the extent of proximity of *gyrB* to *oriC* is a very subjective parameter. After the immediate neighbourhoods of *gyrA* and *gyrB* in the same list of bacteria were assessed, it became evident that while the genetic environment of *gyrA* is not conserved, the genes lying 5' to *gyrB* are frequently conserved. Therefore, more robust grouping was done according to: 1 – operonic or separate arrangements of *gyrA* and *gyrB*, 2 – conservation of the genetic environment 5' of *gyrB* (Table 5.1). The four groups that have been identified represent all the possible arrangements of these genes in bacteria, since bacteria with any of the two genes missing were not detected, and probably do not exist due to the essentiality of DNA gyrase.

Genes that are conserved lying 5' to *gyrB* in bacteria belonging to the groups 1 and 3 are homologues of *dnaA*, *dnaN* and *recF*. These genes are conserved in their orientation – they are aligned in tandem with *gyrB*. Transcription from these co-oriented neighbouring genes provides a strong input of DNA relaxation (Sobetzko, 2016) that stimulates transcription of supercoiling-sensitive P_{gyrB} (Menzel & Gellert, 1987). I conclude that this strong local DNA relaxation obtained from transcription of the neighbouring genes is a driving force for the conservation of *gyrB* position. This and additional evidence, such that the inability to delete *gyrB* from the *gyrAB_gyrB Salmonella* strain during an attempt to construct *gyrAB* (Section 5.2.1); differential expressional behaviour of *gyrA* and *gyrB* in response to quinolones (Neumann & Quiñones, 1997); together with the absolute lack of conservation of *gyrA* position, indicate that the position of *gyrB*, but not that of *gyrA*, is essential. No patterns in the position of *gyrB* when it is not conserved and in the position of *gyrA* were found.

Because it is difficult to conclude whether a certain taxon is enriched in a particular group by looking at table 5.1, a phylogenetic tree was plotted to include all of the bacteria in the table (Fig. 6.2). This tree was constructed using the phylogenetic tree generator phyloT that is based on NCBI taxonomy (Letunic & Bork, 2019) and positions of the branches were manually reviewed with aid of NCBI taxonomy browser. It can be seen that one group can be present in multiple unrelated phyla and similarly, one phylum can contain members of several groups, illustrating a high diversity of *gyrA* and *gyrB* chromosomal arrangements. Certain patterns are, however, visible. Bacteria of the group 1 are exclusively located within phylum Proteobacteria. All the members of phylum Bacteroidetes that were checked belong to the group 2, but other phyla can similarly contain some members of the group 2. There is a high frequency of group 3 arrangement within the superphylum Terrabacteria (Firmicutes, Tenericutes, Actinobacteria, Deinococcus), although this arrangement can be encountered elsewhere too. Finally, all the tree members of the group 4 shown belong to the phylum Chlamydiae. No other phylum was found to contain bacteria of the group 4, but the existence of the group 4 arrangement outside of the Chlamydiae cannot be ruled out. It is also hard to draw parallels between the lifestyle of an organism and the group to which it belongs, since bacteria of various lifestyles can be members of the same group. While *gyrA* and *gyrB* locations in multiple bacteria were scanned (many not shown) every possibility remains that further sampling will broaden the existing groups and reveal further details.

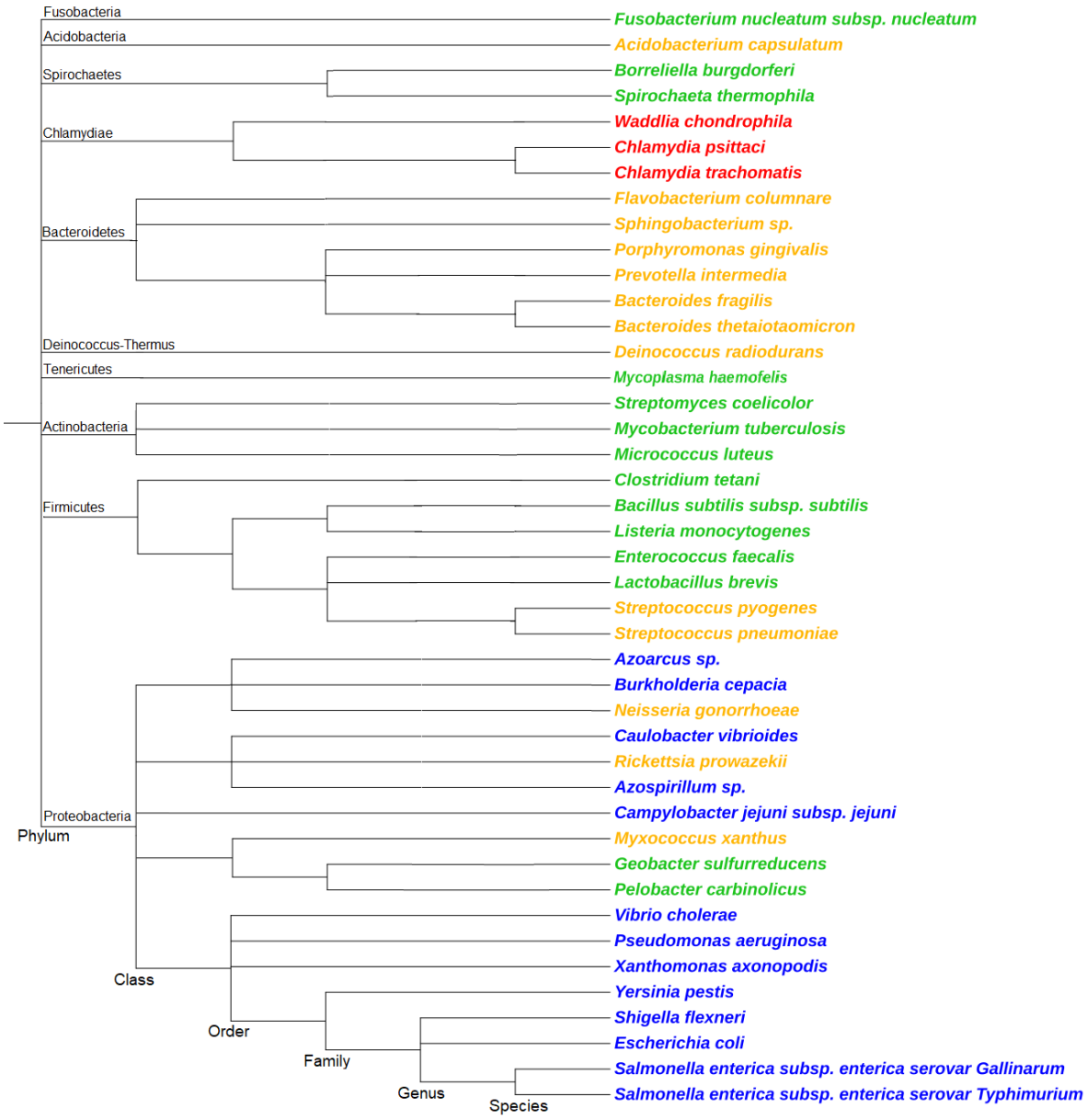


Figure 6.2. Phylogenetic tree of bacteria that belong to different groups of the *gyrA* and *gyrB* arrangement. The phylogenetic tree was built in the phyloT – phylogenetic tree generator based on NCBI taxonomy. The four groups of *gyrA* and *gyrB* arrangements are indicated by colours. Blue – group 1 (separate positions of *gyrA* and *gyrB*, conserved genetic environment 5' from *gyrB*). Orange – group 2 (separate positions of *gyrA* and *gyrB*, non-conserved genetic environment 5' from *gyrB*, Green – group 3 (*gyrBA* operon, conserved genetic environment 5' from *gyrB*). Red – group 4 (*gyrBA* operon, non-conserved genetic environment 5' from *gyrB*). Phyla names are indicated.

6.1.5 Aims of this work revisited

Most of the aims formulated in the section 1.1.8 have been successfully achieved. First, the OrfSwap$^{ihfA\text{-}ihfB}$ – strain that contained the reciprocal ORF exchange of *ihfA* and *ihfB*, as well as the *ihfAB* and *ihfBA* – strains that harboured the corresponding operons were constructed to study the significance of the chromosomal positions of the genes that encode IHF. To study the significance of the separate positioning of the genes that encode DNA gyrase, the strain where these genes are arranged in the *gyrBA* operon was constructed. The construction of the strain with the *gyrAB* operon failed. Although the *gyrAB* operon itself was constructed, it was not possible to delete the original copy of *gyrB,* indicating the importance of its position. The approach to gene repositioning pioneered in this study is creation of artificial operons. All the constructed strains underwent whole genome sequencing to show that they differ from the WT only at the loci of genetic repositioning: they contained no other genetic alterations.

Multiple phenotypic tests of the constructed strains were performed. These included testing of growth dynamics, virulence characteristics and an examination of phenotypes associated with IHF and DNA gyrase. Many phenotypes remained unaltered, highlighting *Salmonella's* ability to tolerate repositioning of its important regulatory genes without displaying significant deleterious effect. The phenotypes that were altered showed that operonic arrangement of the IHF-encoding genes has milder consequences than the reciprocal exchange of these genes. The altered ability of the *gyrBA* strain to supercoil its DNA that can be tuned by changing growth conditions, points at potential differences in global gene expression in this strain.

The *ihfA* and *ihfB* repositionings resulted in rescheduled transcription of these genes that was unique for every type of rearrangement, further highlighting milder consequences of IHF operons compared to the ORF exchange. The operonic arrangement of *gyrA* and *gyrB* did not affect transcription of these genes at all. Interestingly, the measurements of IHF protein production dynamics revealed a greater abundance of IhfB than IhfA in the WT and an exchange in subunit abundance in the OrfSwap$^{ihfA\text{-}ihfB}$ strain, highlighting the

importance of the 5' and 3' regulatory regions for mRNA stability, leading in turn to the unequal production of IHF subunits.

The excess of IhfB over IhfA in the WT and the opposite pattern in the OrfSwap$^{ihfA\text{-}ihfB}$ strain, was proved with mass spectrometry. This is the first study showing that IhfA and IhfB are produced in a ratio other than 1:1 and can be considered in the context of previous studies that hypothesised the possibility of IhfB homodimer formation. Relative *gyrA* and *gyrB* positions in bacteria belonging to various phyla were assessed, allowing the categorisation of the possible arrangements into four groups based on the relative proximity and conservation of genetic environment 5' from *gyrB*. An enrichment of members of particular groups among representatives of particular taxa was observed, but it was not absolute. An importance of the conserved *gyrB* position was revealed, but the selective pressures responsible for *gyrA* positioning and *gyrB* positioning when its genetic environment was not conserved were not elucidated. This comparative study was the first to attempt the categorisation of *gyrA* and *gyrB* positions. Although the reasons behind an acquisition of the operonic versus separate positioning of these genes were not determined, a step forward was made by determining patterns in the *gyrA* and *gyrB* arrangement by breaking down the problem into smaller parts.

There was no aim to study global gene expression in every repositioned strain, instead it was implied that the required study will be performed in the future if deemed necessary. Unfortunately, there was not enough time to carry out such study in the *gyrBA* strain, but it will be a solid step towards elucidating the advantage behind the separate positioning of *gyrA* and *gyrB* in *Salmonella*, if work with the *gyrBA* strain will be continued in the future. The study that compared the proteomes of the WT and the OrfSwap$^{ihfA\text{-}ihfB}$ strain was proposed to solve the IHF stoichiometry question. Alongside, it revealed a number of differentially expressed proteins with a bias towards downregulation in the OrfSwap$^{ihfA\text{-}ihfB}$ strain that was due to decreased expression of ribosomal proteins and, therefore, lower translation capacity.

Chromosomal gene repositionings were confirmed to be a useful tool to study evolutionary forces behind the creation of existing gene arrangements. The

insights gained provide useful guides for future work in the building of artificial chromosomes in synthetic biology. The findings reported in this thesis significantly contribute to our knowledge of the significance of chromosomal positions of genes that encode IHF and DNA gyrase, as well as biology of these proteins.

Bibliography

Adkins, J.N., Mottaz, H.M., Norbeck, A.D., Gustin, J.K., Rue, J., Clauss, T.R., Purvine, S.O., Rodland, K.D., Heffron, F., and Smith, R.D. (2006) Analysis of the *Salmonella* Typhimurium proteome through environmental response toward infectious conditions. *Mol Cell Proteomics* **5**: 1450-1461.

Ali Azam, T., Hiraga, S., and Ishihama, A. (2000) Two types of localization of the DNA-binding proteins within the *Escherichia coli* nucleoid. *Genes Cells* **5**: 613-626.

Ali Azam, T., and Ishihama, A. (2015) Growth phase dependent changes in the structure and protein composition of nucleoid in *Escherichia coli*. *Sci China Life Sci* **58**: 902-911.

Ali Azam, T., Iwata, A., Nishimura, A., Ueda, S., and Ishihama, A. (1999) Growth phase-dependent variation in protein composition of the *Escherichia coli* nucleoid. *J Bacteriol* **181**: 6361-6370.

Almirón, M., Link, A.J., Furlong, D., and Kolter, R. (1992) A novel DNA-binding protein with regulatory and protective roles in starved Escherichia coli. *Genes Dev* **6**: 2646-2654.

Alokam, S., Liu, S.L., Said, K., and Sanderson, K.E. (2002) Inversions over the terminus region in *Salmonella* and *Escherichia coli*: IS200s as the sites of homologous recombination inverting the chromosome of *Salmonella enterica* serovar Typhi. *J Bacteriol* **184**: 6190-6197.

Altuvia, S., Almirón, M., Huisman, G., Kolter, R., and Storz, G. (1994) The *dps* promoter is activated by OxyR during growth and by IHF and sigma S in stationary phase. *Mol Microbiol* **13**: 265-272.

Arike, L., Valgepea, K., Peil, L., Nahku, R., Adamberg, K., and Vilu, R. (2012) Comparison and applications of label-free absolute proteome quantification methods on *Escherichia coli*. *J Proteomics* **75**: 5437-5448.

Ashley, R.E., Dittmore, A., McPherson, S.A., Turnbough, C.L., Jr., Neuman, K.C., and Osheroff, N. (2017) Activities of gyrase and topoisomerase IV on positively supercoiled DNA. *Nucleic Acids Res* **45**: 9611-9624.

Aviv, M., Giladi, H., Schreiber, G., Oppenheim, A.B., and Glaser, G. (1994) Expression of the genes coding for the *Escherichia coli* integration host factor are controlled by growth phase, *rpoS*, ppGpp and by autoregulation. *Mol Microbiol* **14**: 1021-1031.

Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K.A., Tomita, M., Wanner, B.L., and Mori, H. (2006) Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol* **2**: 2006.0008.

Badrinarayanan, A., Le, T.B., and Laub, M.T. (2015) Bacterial chromosome organization and segregation. *Annu Rev Cell Dev Biol* **31**: 171-199.

Badrinarayanan, A., Reyes-Lamothe, R., Uphoff, S., Leake, M.C., and Sherratt, D.J. (2012) In vivo architecture and action of bacterial structural maintenance of chromosome proteins. *Science* **338**: 528-531.

Bajaj, V., Lucas, R.L., Hwang, C., and Lee, C.A. (1996) Co-ordinate regulation of *Salmonella* Typhimurium invasion genes by environmental and regulatory factors is mediated by control of *hilA* expression. *Mol Microbiol* **22**: 703-714.

Balke, V.L., and Gralla, J.D. (1987) Changes in the linking number of supercoiled DNA accompany growth transitions in *Escherichia coli*. *J Bacteriol* **169**: 4499-4506.

Ball, C.A., Osuna, R., Ferguson, K.C., and Johnson, R.C. (1992) Dramatic changes in Fis levels upon nutrient upshift in *Escherichia coli*. *J Bacteriol* **174**: 8043-8056.

Bartley, B.A., Kim, K., Medley, J.K., and Sauro, H.M. (2017) Synthetic Biology: Engineering Living Systems from Biophysical Principles. *Biophys J* **112**: 1050-1058.

Bashor, C.J., Horwitz, A.A., Peisajovich, S.G., and Lim, W.A. (2010) Rewiring cells: synthetic biology as a tool to interrogate the organizational principles of living systems. *Annu Rev Biophys* **39**: 515-537.

Bates, A.D., and Maxwell, A. (2007) Energy coupling in type II topoisomerases: why do they hydrolyze ATP? *Biochemistry* **46**: 7929-7941.

Bickle, T.A., and Krüger, D.H. (1993) Biology of DNA restriction. *Microbiol Rev* **57**: 434-450.
Blattner, F.R., Plunkett, G., 3rd, Bloch, C.A., Perna, N.T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J.D., Rode, C.K., Mayhew, G.F., Gregor, J., Davis, N.W., Kirkpatrick, H.A., Goeden, M.A., Rose, D.J., Mau, B., and Shao, Y. (1997) The complete genome sequence of *Escherichia coli* K-12. *Science* **277**: 1453-1462.
Bobay, L.M., and Ochman, H. (2017) The evolution of bacterial genome architecture. *Front Genet* **8**: 72.
Bogue, M.M., Mogre, A., Beckett, M.C., Thomson, N.R., and Dorman, C.J. (2020) Network rewiring: physiological consequences of reciprocally exchanging the physical locations and growth-phase-dependent expression patterns of the *Salmonella fis* and *dps* genes. *mBio* **11**(5): e0218-20.
Bonifield, H.R., and Hughes, K.T. (2003) Flagellar phase variation in *Salmonella enterica* is mediated by a posttranscriptional control mechanism. *J Bacteriol* **185**: 3567-3574.
Booker, B.M., Deng, S., and Higgins, N.P. (2010) DNA topology of highly transcribed operons in *Salmonella enterica* serovar Typhimurium. *Mol Microbiol* **78**: 1348-1364.
Brambilla, E., and Sclavi, B. (2015) Gene regulation by H-NS as a function of growth conditions depends on chromosomal position in *Escherichia coli*. *G3 (Bethesda)* **5**: 605-614.
Brenner, F.W., Villar, R.G., Angulo, F.J., Tauxe, R., and Swaminathan, B. (2000) *Salmonella* nomenclature. *J Clin Microbiol* **38**: 2465-2467.
Browning, D.F., Cole, J.A., and Busby, S.J. (2000) Suppression of FNR-dependent transcription activation at the *Escherichia coli nir* promoter by Fis, IHF and H-NS: modulation of transcription initiation by a complex nucleo-protein assembly. *Mol Microbiol* **37**: 1258-1269.
Broz, P., Ohlson, M.B., and Monack, D.M. (2012) Innate immune response to *Salmonella* Typhimurium, a model enteric pathogen. *Gut Microbes* **3**: 62-70.
Bryant, J.A., Sellars, L.E., Busby, S.J., and Lee, D.J. (2014) Chromosome position effects on gene expression in *Escherichia coli* K-12. *Nucleic Acids Res* **42**: 11383-11392.
Bullas, L.R., and Ryu, J.I. (1983) *Salmonella* Typhimurium LT2 strains which are r- m+ for all three chromosomally located systems of DNA restriction and modification. *J Bacteriol* **156**: 471-474.
Bush, N.G., Evans-Roberts, K., and Maxwell, A. (2015) DNA Topoisomerases. *EcoSal Plus* **6**.
Bushman, W., Yin, S., Thio, L.L., and Landy, A. (1984) Determinants of directionality in lambda site-specific recombination. *Cell* **39**: 699-706
Cabrera, J.E., Cagliero, C., Quan, S., Squires, C.L., and Jin, D.J. (2009) Active transcription of rRNA operons condenses the nucleoid in *Escherichia coli*: examining the effect of transcription on nucleoid structure in the absence of transertion. *J Bacteriol* **191**: 4180-4185.
Cameron, A.D., and Dorman, C.J. (2012) A fundamental regulatory mechanism operating through OmpR and DNA topology controls expression of *Salmonella* pathogenicity islands SPI-1 and SPI-2. *PLoS Genet* **8**: e1002615.
Cameron, A.D., Stoebel, D.M., and Dorman, C.J. (2011) DNA supercoiling is differentially regulated by environmental factors and FIS in *Escherichia coli* and *Salmonella enterica*. *Mol Microbiol* **80**: 85-101.
Cameron, A.D.S., Dillon, S.C., Kröger, C., Beran, L., and Dorman, C.J. (2017) Broad-scale redistribution of mRNA abundance and transcriptional machinery in response to growth rate in *Salmonella enterica* serovar Typhimurium. *Microb Genom* **3**: e000127.
Campbell, E.A., Korzheva, N., Mustaev, A., Murakami, K., Nair, S., Goldfarb, A., and Darst, S.A. (2001) Structural mechanism for rifampicin inhibition of bacterial rna polymerase. *Cell* **104**: 901-912.
Cello, J., Paul, A.V., and Wimmer, E. (2002) Chemical synthesis of poliovirus cDNA: generation of infectious virus in the absence of natural template. *Science* **297**: 1016-1018.

Chaban, B., Hughes, H.V., and Beeby, M. (2015) The flagellum in bacterial pathogens: For motility and a whole lot more. *Semin Cell Dev Biol* **46**: 91-103.
Chaconas, G., and Kobryn, K. (2010) Structure, function, and evolution of linear replicons in *Borrelia*. *Annu Rev Microbiol* **64**: 185-202.
Chakraborty, S., Mizusaki, H., and Kenney, L.J. (2015) A FRET-based DNA biosensor tracks OmpR-dependent acidification of *Salmonella* during macrophage infection. *PLoS Biol* **13**: e1002116.
Champoux, J.J. (2001) DNA topoisomerases: structure, function, and mechanism. *Annu Rev Biochem* **70**: 369-413.
Chen, C.Y., Ezzeddine, N., and Shyu, A.B. (2008) Messenger RNA half-life measurements in mammalian cells. *Methods Enzymol* **448**: 335-357.
Cherepanov, P.P., and Wackernagel, W. (1995) Gene disruption in *Escherichia coli*: TcR and KmR cassettes with the option of Flp-catalyzed excision of the antibiotic-resistance determinant. *Gene* **158**: 9-14.
Cho, B.K., Barrett, C.L., Knight, E.M., Park, Y.S., and Palsson, B. (2008) Genome-scale reconstruction of the Lrp regulatory network in *Escherichia coli*. *Proc Natl Acad Sci U S A* **105**: 19462-19467.
Cho, B.K., Knight, E.M., and Palsson, B.O. (2006) PCR-based tandem epitope tagging system for *Escherichia coli* genome engineering. *Biotechniques* **40**: 67-72.
Chong, S., Chen, C., Ge, H., and Xie, X.S. (2014) Mechanism of transcriptional bursting in bacteria. *Cell* **158**: 314-326.
Cirillo, D.M., Valdivia, R.H., Monack, D.M., and Falkow, S. (1998) Macrophage-dependent induction of the *Salmonella* pathogenicity island 2 type III secretion system and its role in intracellular survival. *Mol Microbiol* **30**: 175-188.
Clark, L., Perrett, C.A., Malt, L., Harward, C., Humphrey, S., Jepson, K.A., Martinez-Argudo, I., Carney, L.J., La Ragione, R.M., Humphrey, T.J., and Jepson, M.A. (2011) Differences in Salmonella enterica serovar Typhimurium strain invasiveness are associated with heterogeneity in SPI-1 gene expression. *Microbiology (Reading)* **157**: 2072-2083.
Claret, L., and Rouviere-Yaniv, J. (1997) Variation in HU composition during growth of *Escherichia coli*: the heterodimer is required for long term survival. *J Mol Biol* **273**: 93-104.
Colgan, A.M., Quinn, H.J., Kary, S.C., Mitchenall, L.A., Maxwell, A., Cameron, A.D.S., and Dorman, C.J. (2018) Negative supercoiling of DNA by gyrase is inhibited in *Salmonella enterica* serovar Typhimurium during adaptation to acid stress. *Mol Microbiol* **107**: 734-746.
Colland, F., Barth, M., Hengge-Aronis, R., and Kolb, A. (2000) sigma factor selectivity of *Escherichia coli* RNA polymerase: role for CRP, IHF and lrp transcription factors. *EMBO J* **19**: 3028-3037.
Connolly, M., Arra, A., Zvoda, V., Steinbach, P.J., Rice, P.A., and Ansari, A. (2018) Static kinks or flexible hinges: multiple conformations of Bent DNA bound to Integration Host Factor revealed by fluorescence lifetime measurements. *J Phys Chem B* **122**: 11519-11534.
Cooper, S., and Helmstetter, C.E. (1968) Chromosome replication and the division cycle of *Escherichia coli* B/r. *J Mol Biol* **31**: 519-540.
Cornish, R.M., Roth, J.R., and Poulter, C.D. (2006) Lethal mutations in the isoprenoid pathway of Salmonella enterica. *J Bacteriol* **188**: 1444-1450.
Correia, S., Nunes-Miranda, J.D., Pinto, L., Santos, H.M., de Toro, M., Sáenz, Y., Torres, C., Capelo, J.L., Poeta, P., and Igrejas, G. (2014) Complete proteome of a quinolone-resistant Salmonella Typhimurium phage type DT104B clinical strain. *Int J Mol Sci* **15**: 14191-14219.

Cosgriff, S., Chintakayala, K., Chim, Y.T., Chen, X., Allen, S., Lovering, A.L., and Grainger, D.C. (2010) Dimerization and DNA-dependent aggregation of the Escherichia coli nucleoid protein and chaperone CbpA. *Mol Microbiol* **77**: 1289-1300.

Cox, J., Hein, M.Y., Luber, C.A., Paron, I., Nagaraj, N., and Mann, M. (2014) Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol Cell Proteomics* **13**: 2513-2526.

Crellin, P., Sewitz, S., and Chalmers, R. (2004) DNA looping and catalysis; the IHF-folded arm of Tn10 promotes conformational changes and hairpin resolution. *Mol Cell* **13**: 537-547.

Crisona, N.J., and Cozzarelli, N.R. (2006) Alteration of *Escherichia coli* topoisomerase IV conformation upon enzyme binding to positively supercoiled DNA. *J Biol Chem* **281**: 18927-18932.

Cui, T., Moro-oka, N., Ohsumi, K., Kodama, K., Ohshima, T., Ogasawara, N., Mori, H., Wanner, B., Niki, H., and Horiuchi, T. (2007) *Escherichia coli* with a linear genome. *EMBO Rep* **8**: 181-187.

Dame, R.T., Kalmykowa, O.J., and Grainger, D.C. (2011) Chromosomal macrodomains and associated proteins: implications for DNA organization and replication in gram negative bacteria. *PLoS Genet* **7**: e1002123.

Dame, R.T., Noom, M.C., and Wuite, G.J. (2006) Bacterial chromatin organization by H-NS protein unravelled using dual DNA manipulation. *Nature* **444**: 387-390.

Dame, R.T., Rashid, F.M., and Grainger, D.C. (2020) Chromosome organization in bacteria: mechanistic insights into genome structure and function. *Nat Rev Genet* **21**: 227-242.

Danilova, O., Reyes-Lamothe, R., Pinskaya, M., Sherratt, D., and Possoz, C. (2007) MukB colocalizes with the *oriC* region and is required for organization of the two *Escherichia coli* chromosome arms into separate cell halves. *Mol Microbiol* **65**: 1485-1492.

Datsenko, K.A., and Wanner, B.L. (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci U S A* **97**: 6640-6645.

Deatherage, D.E., and Barrick, J.E. (2014) Identification of mutations in laboratory-evolved microbes from next-generation sequencing data using breseq. *Methods Mol Biol* **1151**: 165-188.

de los Rios, S., and Perona, J.J. (2007) Structure of the Escherichia coli leucine-responsive regulatory protein Lrp reveals a novel octameric assembly. *J Mol Biol* **366**: 1589-1602.

Dey, D., Nagaraja, V., and Ramakumar, S. (2017) Structural and evolutionary analyses reveal determinants of DNA binding specificities of nucleoid-associated proteins HU and IHF. *Mol Phylogenet Evol* **107**: 356-366.

Dillon, S.C., and Dorman, C.J. (2010) Bacterial nucleoid-associated proteins, nucleoid structure and gene expression. *Nat Rev Microbiol* **8**: 185-195.

Dilweg, I.W., and Dame, R.T. (2018) Post-translational modification of nucleoid-associated proteins: an extra layer of functional modulation in bacteria? *Biochem Soc Trans* **46**: 1381-1392.

DiNardo, S., Voelkel, K.A., Sternglanz, R., Reynolds, A.E., and Wright, A. (1982) *Escherichia coli* DNA topoisomerase I mutants have compensatory mutations in DNA gyrase genes. *Cell* **31**: 43-51.

Ditto, M.D., Roberts, D., and Weisberg, R.A. (1994) Growth phase variation of integration host factor level in *Escherichia coli*. *J Bacteriol* **176**: 3738-3748.

Doetsch, M., Gstrein, T., Schroeder, R., and Fürtig, B. (2010) Mechanisms of StpA-mediated RNA remodeling. *RNA Biol* **7**: 735-743.

Doolittle, R.F., Feng, D.F., Tsang, S., Cho, G., and Little, E. (1996) Determining divergence times of the major kingdoms of living organisms with a protein clock. *Science* **271**: 470-477.

Dorman, C.J., (2013) Genome architecture and global gene regulation in bacteria: making progress towards a unified model? *Nat Rev Microbiol.* 11: 349-355.

Dorman, C.J. (2014) Function of nucleoid-associated proteins in chromosome structuring and transcriptional regulation. *J Mol Microbiol Biotechnol* **24**: 316-331.
Dorman, C.J., (2020) *Structure and Function of the Bacterial Genome*. John Wiley and Sons, Hoboken, NJ.
Dorman, C.J., and Bogue, M.M. (2016) The interplay between DNA topology and accessory factors in site-specific recombination in bacteria and their bacteriophages. *Sci Prog* **99**: 420-437.
Dorman, C.J., and Dorman, M.J. (2016) DNA supercoiling is a fundamental regulatory principle in the control of bacterial gene expression. *Biophys Rev* **8**: 209-220.
Drecktrah, D., Knodler, L.A., Ireland, R., and Steele-Mortimer, O. (2006) The mechanism of *Salmonella* entry determines the vacuolar environment and intracellular gene expression. *Traffic* **7**: 39-51.
Drlica, K., and Snyder, M. (1978) Superhelical *Escherichia coli* DNA: relaxation by coumermycin. *J Mol Biol* **120**: 145-154.
Drlica, K., and Zhao, X. (1997) DNA gyrase, topoisomerase IV, and the 4-quinolones. *Microbiol Mol Biol Rev* **61**: 377-392.
Eichler, E.E., and Sankoff, D. (2003) Structural dynamics of eukaryotic chromosome evolution. *Science* **301**: 793-797.
Ellis, T., Wang, X., and Collins, J.J. (2009) Diversity-based, model-guided construction of synthetic gene networks with predicted functions. *Nat Biotechnol* **27**: 465-471.
Espeli, O., Levine, C., Hassing, H., and Marians, K.J. (2003) Temporal regulation of topoisomerase IV activity in E. coli. *Mol Cell* **11**: 189-201.
Fass, E., and Groisman, E.A. (2009) Control of Salmonella pathogenicity island-2 gene expression. *Curr Opin Microbiol* **12**: 199-204.
Fei, J., and Sharma, C.M. (2018) RNA Localization in Bacteria. *Microbiol Spectr* **6**.
Ferrándiz, M.J., Martín-Galiano, A.J., Schvartzman, J.B., and de la Campa, A.G. (2010) The genome of *Streptococcus pneumoniae* is organized in topology-reacting gene clusters. *Nucleic Acids Res* **38**: 3570-3581.
Figueira, R., and Holden, D.W. (2012) Functions of the *Salmonella* pathogenicity island 2 (SPI-2) type III secretion system effectors. *Microbiology (Reading)* **158**: 1147-1161.
Finlay, B.B., and Brumell, J.H. (2000) *Salmonella* interactions with host cells: in vitro to in vivo. *Philos Trans R Soc Lond B Biol Sci* **355**: 623-631.
Fitzgerald, D.M., Bonocora, R.P., and Wade, J.T. (2014) Comprehensive mapping of the *Escherichia coli* flagellar regulatory network. *PLoS Genet* **10**: e1004649.
Fitzgerald, S., Dillon, S.C., Chao, T.C., Wiencko, H.L., Hokamp, K., Cameron, A.D., and Dorman, C.J. (2015) Re-engineering cellular physiology by rewiring high-level global regulatory genes. *Sci Rep* **5**: 17653.
Frey, B., and Suppmann, B. (1995) Demonstration of the Expand PCR system's greater fidelity and higher yields with a lacI-based PCR fidelity assay. *Biochemica* **2**: 8-9.
Galán, J.E. (1996) Molecular genetic bases of *Salmonella* entry into host cells. *Mol Microbiol* **20**: 263-271.
Gerganova, V., Berger, M., Zaldastanishvili, E., Sobetzko, P., Lafon, C., Mourez, M., Travers, A., and Muskhelishvili, G. (2015) Chromosomal position shift of a regulatory gene alters the bacterial phenotype. *Nucleic Acids Res* **43**: 8215-8226.
Giai Gianetto, Q., Couté, Y., Bruley, C., and Burger, T. (2016) Uses and misuses of the fudge factor in quantitative discovery proteomics. *Proteomics* **16**: 1955-1960.
Gibson, D.G., Benders, G.A., Andrews-Pfannkoch, C., Denisova, E.A., Baden-Tillson, H., Zaveri, J., Stockwell, T.B., Brownley, A., Thomas, D.W., Algire, M.A., Merryman, C., Young, L., Noskov, V.N., Glass, J.I., Venter, J.C., Hutchison, C.A., 3rd, and Smith, H.O. (2008) Complete chemical synthesis, assembly, and cloning of a *Mycoplasma genitalium* genome. *Science* **319**: 1215-1220.

Gilchrist, J.J., and MacLennan, C.A. (2019) Invasive Nontyphoidal *Salmonella* Disease in Africa. *EcoSal Plus* **8**.

Glass, J.I., Assad-Garcia, N., Alperovich, N., Yooseph, S., Lewis, M.R., Maruf, M., Hutchison, C.A., 3rd, Smith, H.O., and Venter, J.C. (2006) Essential genes of a minimal bacterium. *Proc Natl Acad Sci U S A* **103**: 425-430.

Glöckner, F.O., Yilmaz, P., Quast, C., Gerken, J., Beccati, A., Ciuprina, A., Bruns, G., Yarza, P., Peplies, J., Westram, R., and Ludwig, W. (2017) 25 years of serving the community with ribosomal RNA gene reference databases and tools. *J Biotechnol* **261**: 169-176.

Goodrich, J.A., Schwartz, M.L., and McClure, W.R. (1990) Searching for and predicting the activity of sites for DNA binding proteins: compilation and analysis of the binding sites for *Escherichia coli* integration host factor (IHF). *Nucleic Acids Res* **18**: 4993-5000.

Gordon, B.R., Li, Y., Cote, A., Weirauch, M.T., Ding, P., Hughes, T.R., Navarre, W.W., Xia, B., and Liu, J. (2011) Structural basis for recognition of AT-rich DNA by unrelated xenogeneic silencing proteins. *Proc Natl Acad Sci U S A* **108**: 10690-10695.

Grimwade, J.E., Ryan, V.T., and Leonard, A.C. (2000) IHF redistributes bound initiator protein, DnaA, on supercoiled *oriC* of *Escherichia coli*. *Mol Microbiol* **35**: 835-844.

Grove, A. (2011) Functional evolution of bacterial histone-like HU proteins. *Curr Issues Mol Biol* **13**: 1-12.

Guo, F., and Adhya, S. (2007) Spiral structure of Escherichia coli HUalphabeta provides foundation for DNA supercoiling. *Proc Natl Acad Sci U S A* **104**: 4309-4314.

Gyorfy, Z., Draskovits, G., Vernyik, V., Blattner, F.F., Gaal, T., and Posfai, G. (2015) Engineered ribosomal RNA operon copy-number variants of *E. coli* reveal the evolutionary trade-offs shaping rRNA operon number. *Nucleic Acids Res* **43**: 1783-1794.

Haran, T.E., and Mohanty, U. (2009) The unique structure of A-tracts and intrinsic DNA bending. *Q Rev Biophys* **42**: 41-81.

Hardy, C.D., and Cozzarelli, N.R. (2005) A genetic selection for supercoiling mutants of *Escherichia coli* reveals proteins implicated in chromosome structure. *Mol Microbiol* **57**: 1636-1652.

Hartman, P.E., Loper, J.C., and Serman, D. (1960) Fine structure mapping by complete transduction between histidine-requiring *Salmonella* mutants. *J Gen Microbiol* **22**: 323-353.

Hautefort, I., Proenca, M.J., and Hinton, J.C. (2003) Single-copy green fluorescent protein gene fusions allow accurate measurement of Salmonella gene expression in vitro and during infection of mammalian cells. *Appl Environ Microbiol* **69**: 7480-7491.

Heather, J.M., and Chain, B. (2016) The sequence of sequencers: The history of sequencing DNA. *Genomics* **107**: 1-8.

Heidelberg, J.F., Eisen, J.A., Nelson, W.C., Clayton, R.A., Gwinn, M.L., Dodson, R.J., Haft, D.H., Hickey, E.K., Peterson, J.D., Umayam, L., Gill, S.R., Nelson, K.E., Read, T.D., Tettelin, H., Richardson, D., Ermolaeva, M.D., Vamathevan, J., Bass, S., Qin, H., Dragoi, I., Sellers, P., McDonald, L., Utterback, T., Fleishmann, R.D., Nierman, W.C., White, O., Salzberg, S.L., Smith, H.O., Colwell, R.R., Mekalanos, J.J., Venter, J.C., and Fraser, C.M. (2000) DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*. *Nature* **406**: 477-483.

Henneman, B., Heinsman, J., Battjes, J., and Dame, R.T. (2018) Quantitation of DNA-binding affinity using tethered particle motion. *Methods Mol Biol* **1837**: 257-275.

Hicks, M., Bachmann, T.T., and Wang, B. (2020) Synthetic biology enables programmable cell-based biosensors. *Chemphyschem* **21**: 132-144.

Higgins, N.P. (2014) RNA polymerase: chromosome domain boundary maker and regulator of supercoil density. *Curr Opin Microbiol* **22**: 138-143.

Higgins, N.P., Peebles, C.L., Sugino, A., and Cozzarelli, N.R. (1978) Purification of subunits of Escherichia coli DNA gyrase and reconstitution of enzymatic activity. *Proc Natl Acad Sci U S A* **75**: 1773-1777.

Higgins, N.P., and Vologodskii, A.V. (2015) Topological Behavior of Plasmid DNA. *Microbiol Spectr* **3**.

Hoiseth, S.K., and Stocker, B.A. (1981) Aromatic-dependent *Salmonella* Typhimurium are non-virulent and effective as live vaccines. *Nature* **291**: 238-239.

Huang, T.W., Hsu, C.C., Yang, H.Y., and Chen, C.W. (2013) Topoisomerase IV is required for partitioning of circular chromosomes but not linear chromosomes in *Streptomyces*. *Nucleic Acids Res* **41**: 10403-10413.

Hui, M.P., Foley, P.L., and Belasco, J.G. (2014) Messenger RNA degradation in bacterial cells. *Annu Rev Genet* **48**: 537-559.

Hutchison, C.A., 3rd, Chuang, R.Y., Noskov, V.N., Assad-Garcia, N., Deerinck, T.J., Ellisman, M.H., Gill, J., Kannan, K., Karas, B.J., Ma, L., Pelletier, J.F., Qi, Z.Q., Richter, R.A., Strychalski, E.A., Sun, L., Suzuki, Y., Tsvetanova, B., Wise, K.S., Smith, H.O., Glass, J.I., Merryman, C., Gibson, D.G., and Venter, J.C. (2016) Design and synthesis of a minimal bacterial genome. *Science* **351**: aad6253.

Ibarra, J.A., Knodler, L.A., Sturdevant, D.E., Virtaneva, K., Carmody, A.B., Fischer, E.R., Porcella, S.F., and Steele-Mortimer, O. (2010) Induction of *Salmonella* pathogenicity island 1 under different growth conditions can affect *Salmonella*-host cell interactions in vitro. *Microbiology* **156**: 1120-1133.

Ilyas, B., Tsai, C.N., and Coombes, B.K. (2017) Evolution of *Salmonella*-host cell interactions through a dynamic bacterial genome. *Front Cell Infect Microbiol* **7**: 428.

Jacob, F., and Monod, J. (1961) Genetic regulatory mechanisms in the synthesis of proteins. *J Mol Biol* **3**: 318-356.

Jain, P., and Nagaraja, V. (2005) An atypical type II topoisomerase from *Mycobacterium smegmatis* with positive supercoiling activity. *Mol Microbiol* **58**: 1392-1405.

Janissen, R., Arens, M.M.A., Vtyurina, N.N., Rivai, Z., Sunday, N.D., Eslami-Mossallam, B., Gritsenko, A.A., Laan, L., de Ridder, D., Artsimovitch, I., Dekker, N.H., Abbondanzieri, E.A., and Meyer, A.S. (2018) Global DNA Compaction in stationary-phase bacteria does not affect transcription. *Cell* **174**: 1188-1199.e1114.

Jarvik, J.W., and Telmer, C.A. (1998) Epitope tagging. *Annu Rev Genet* **32**: 601-618.

Jeong, K.S., Ahn, J., and Khodursky, A.B. (2004) Spatial patterns of transcriptional activity in the chromosome of *Escherichia coli*. *Genome Biol* **5**: R86.

Kaczanowska, M., and Rydén-Aulin, M. (2007) Ribosome biogenesis and the translation process in *Escherichia coli*. *Microbiol Mol Biol Rev* **71**: 477-494.

Kakeda, M., Ueguchi, C., Yamada, H., and Mizuno, T. (1995) An Escherichia coli curved DNA-binding protein whose expression is affected by the stationary phase-specific sigma factor sigma S. *Mol Gen Genet* **248**: 629-634.

Kamashev, D., Agapova, Y., Rastorguev, S., Talyzina, A.A., Boyko, K.M., Korzhenevskiy, D.A., Vlaskina, A., Vasilov, R., Timofeev, V.I., and Rakitina, T.V. (2017) Comparison of histone-like HU protein DNA-binding properties and HU/IHF protein sequence alignment. *PLoS One* **12**: e0188037.

Kaneko, M., Emoto, Y., and Emoto, M. (2016) A simple, reproducible, inexpensive, yet old-fashioned method for determining phagocytic and bactericidal activities of macrophages. *Yonsei Med J* **57**: 283-290.

Kasho, K., Fujimitsu, K., Matoba, T., Oshima, T., and Katayama, T. (2014) Timely binding of IHF and Fis to DARS2 regulates ATP-DnaA production and replication initiation. *Nucleic Acids Res* **42**: 13134-13149.

Kasho, K., Tanaka, H., Sakai, R., and Katayama, T. (2017) Cooperative DnaA Binding to the Negatively Supercoiled datA Locus Stimulates DnaA-ATP Hydrolysis. *J Biol Chem* **292**: 1251-1266.

Keane, O.M., and Dorman, C.J. (2003) The *gyr* genes of *Salmonella enterica* serovar Typhimurium are repressed by the factor for inversion stimulation, Fis. *Mol Genet Genomics* **270**: 56-65.

Khemici, V., Poljak, L., Luisi, B.F., and Carpousis, A.J. (2008) The RNase E of *Escherichia coli* is a membrane-binding protein. *Mol Microbiol* **70**: 799-813.

Knodler, L.A., Vallance, B.A., Celli, J., Winfree, S., Hansen, B., Montero, M., and Steele-Mortimer, O. (2010) Dissemination of invasive *Salmonella* via bacterial-induced extrusion of mucosal epithelia. *Proc Natl Acad Sci U S A* **107**: 17733-17738.

Kobryn, K., Watson, M.A., Allison, R.G., and Chaconas, G. (2002) The Mu three-site synapse: a strained assembly platform in which delivery of the L1 transposase binding site triggers catalytic commitment. *Mol Cell* **10**: 659-669.

Koster, D.A., Crut, A., Shuman, S., Bjornsti, M.A., and Dekker, N.H. (2010) Cellular strategies for regulating DNA supercoiling: a single-molecule perspective. *Cell* **142**: 519-530.

Kraemer, J.A., Sanderlin, A.G., and Laub, M.T. (2019) The stringent response inhibits DNA replication initiation in *E. coli* by modulating supercoiling of. *mBio* **10**.

Kröger, C., Colgan, A., Srikumar, S., Handler, K., Sivasankaran, S.K., Hammarlof, D.L., Canals, R., Grissom, J.E., Conway, T., Hokamp, K., and Hinton, J.C. (2013) An infection-relevant transcriptomic compendium for *Salmonella enterica* Serovar Typhimurium. *Cell Host Microbe* **14**: 683-695.

Kröger, C., Dillon, S.C., Cameron, A.D., Papenfort, K., Sivasankaran, S.K., Hokamp, K., Chao, Y., Sittka, A., Hébrard, M., Händler, K., Colgan, A., Leekitcharoenphon, P., Langridge, G.C., Lohan, A.J., Loftus, B., Lucchini, S., Ussery, D.W., Dorman, C.J., Thomson, N.R., Vogel, J., and Hinton, J.C. (2012) The transcriptional landscape and small RNAs of *Salmonella enterica* serovar Typhimurium. *Proc Natl Acad Sci U S A* **109**: E1277-1286.

Kurtz, J.R., Goggins, J.A., and McLachlan, J.B. (2017) *Salmonella* infection: Interplay between the bacteria and host immune system. *Immunol Lett* **190**: 42-50.

Laalami, S., Zig, L., and Putzer, H. (2014) Initiation of mRNA decay in bacteria. *Cell Mol Life Sci* **71**: 1799-1828.

Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**: 680-685.

Lago, M., Monteil, V., Douche, T., Guglielmini, J., Criscuolo, A., Maufrais, C., Matondo, M., and Norel, F. (2017) Proteome remodelling by the stress sigma factor RpoS/σ. *Sci Rep* **7**: 2127.

Lal, A., Dhar, A., Trostel, A., Kouzine, F., Seshasayee, A.S., and Adhya, S. (2016) Genome scale patterns of supercoiling in a bacterial chromosome. *Nat Commun* **7**: 11055.

Lawrence, J.G. (1997) Selfish operons and speciation by gene transfer. *Trends Microbiol* **5**: 355-359.

Lawrence, J.G. (2003) Gene organization: selection, selfishness, and serendipity. *Annu Rev Microbiol* **57**: 419-440.

Lawrence, J.G., and Ochman, H. (1997) Amelioration of bacterial genomes: rates of change and exchange. *J Mol Evol* **44**: 383-397.

Lawrence, J.G., and Roth, J.R. (1996) Selfish operons: horizontal transfer may drive the evolution of gene clusters. *Genetics* **143**: 1843-1860.

Le, T.B., Imakaev, M.V., Mirny, L.A., and Laub, M.T. (2013) High-resolution mapping of the spatial organization of a bacterial chromosome. *Science* **342**: 731-734.

Leclerc, G.J., Leclerc, G.M., and Barredo, J.C. (2002) Real-time RT-PCR analysis of mRNA decay: half-life of Beta-actin mRNA in human leukemia CCRF-CEM and Nalm-6 cell lines. *Cancer Cell Int* **2**: 1.

Lee, A.K., Detweiler, C.S., and Falkow, S. (2000) OmpR regulates the two-component system SsrA-SsrB in Salmonella pathogenicity island 2. *J Bacteriol* **182**: 771-781.
Lee, C.A., and Falkow, S. (1990) The ability of *Salmonella* to enter mammalian cells is affected by bacterial growth state. *Proc Natl Acad Sci U S A* **87**: 4304-4308.
Lemire, S., Yehl, K.M., and Lu, T.K. (2018) Phage-based applications in synthetic biology. *Annu Rev Virol* **5**: 453-476.
Lemke, J.J., Sanchez-Vazquez, P., Burgos, H.L., Hedberg, G., Ross, W., and Gourse, R.L. (2011) Direct regulation of *Escherichia coli* ribosomal protein promoters by the transcription factors ppGpp and DksA. *Proc Natl Acad Sci U S A* **108**: 5712-5717.
Letunic, I., and Bork, P. (2019) Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res* **47**: W256-w259.
Lewis, R.J., Singh, O.M., Smith, C.V., Skarzynski, T., Maxwell, A., Wonacott, A.J., and Wigley, D.B. (1996) The nature of inhibition of DNA gyrase by the coumarins and the cyclothialidines revealed by X-ray crystallography. *Embo j* **15**: 1412-1420.
Lin, J., Chen, H., Droge, P., and Yan, J. (2012) Physical organization of DNA by multiple non-specific DNA-binding modes of integration host factor (IHF). *PLoS One* **7**: e49885.
Lipscomb, G.L., Hahn, E.M., Crowley, A.T., and Adams, M.W.W. (2017) Reverse gyrase is essential for microbial growth at 95 °C. *Extremophiles* **21**: 603-608.
Liu, B., and Alberts, B.M. (1995) Head-on collision between a DNA replication apparatus and RNA polymerase transcription complex. *Science* **267**: 1131-1137.
Liu, G., Ma, Q., and Xu, Y. (2018) Physical properties of DNA may direct the binding of nucleoid-associated proteins along the *E. coli* genome. *Math Biosci* **301**: 50-58.
Liu, L.F., and Wang, J.C. (1987) Supercoiling of the DNA template during transcription. *Proc Natl Acad Sci U S A* **84**: 7024-7027.
Liu, Q., Yi, J., Liang, K., Hu, B., Zhang, X., Curtiss, R., and Kong, Q. (2016) Outer membrane vesicles from flagellin-deficient *Salmonella enterica* serovar Typhimurium induce cross-reactive immunity and provide cross-protection against heterologous *Salmonella* challenge. *Sci Rep* **6**: 34776.
Lovett, S.T. (2011) The DNA Exonucleases of *Escherichia coli*. *EcoSal Plus* **4**.
Lu, M., Campbell, J.L., Boye, E., and Kleckner, N. (1994) SeqA: a negative modulator of replication initiation in *E. coli*. *Cell* **77**: 413-426.
Luciano, D.J., Vasilyev, N., Richards, J., Serganov, A., and Belasco, J.G. (2017) A novel RNA phosphorylation state enables 5' End-dependent degradation in *Escherichia coli*. *Mol Cell* **67**: 44-54.e46.
Luo, H., and Gao, F. (2019) DoriC 10.0: an updated database of replication origins in prokaryotic genomes including chromosomes and plasmids. *Nucleic Acids Res* **47**: D74-D77.
Lévi-Meyrueis, C., Monteil, V., Sismeiro, O., Dillies, M.A., Monot, M., Jagla, B., Coppée, J.Y., Dupuy, B., and Norel, F. (2014) Expanding the RpoS/σS-network by RNA sequencing and identification of σS-controlled small RNAs in *Salmonella*. *PLoS One* **9**: e96918.
Macchi, R., Montesissa, L., Murakami, K., Ishihama, A., De Lorenzo, V., and Bertoni, G. (2003) Recruitment of sigma54-RNA polymerase to the *Pu* promoter of *Pseudomonas putida* through integration host factor-mediated positioning switch of alpha subunit carboxyl-terminal domain on an UP-like element. *J Biol Chem* **278**: 27695-27702.
Maciag, A., Peano, C., Pietrelli, A., Egli, T., De Bellis, G., and Landini, P. (2011) In vitro transcription profiling of the σS subunit of bacterial RNA polymerase: re-definition of the σS regulon and identification of σS-specific promoter sequence elements. *Nucleic Acids Res* **39**: 5338-5355.
Macián, F., Pérez-Roger, I., and Armengod, M.E. (1994) An improved vector system for constructing transcriptional *lacZ* fusions: analysis of regulation of the *dnaA*, *dnaN*, *recF* and *gyrB* genes of *Escherichia coli*. *Gene* **145**: 17-24.

Mangan, M.W., Lucchini, S., Danino, V., O'Croinin, T., Hinton, J.C., and Dorman, C.J. (2006) The integration host factor (IHF) integrates stationary-phase and virulence gene expression in *Salmonella enterica* serovar Typhimurium. *Mol Microbiol* **59**: 1831-1847.

Mangan, M.W., Lucchini, S., Ó'Cróinín, T., Fitzgerald, S., Hinton, J.C.D., and Dorman, C.J. (2011) Nucleoid-associated protein HU controls three regulons that coordinate virulence, response to stress and general physiology in *Salmonella enterica* serovar Typhimurium. *Microbiology (Reading)* **157**: 1075-1087.

Manjunatha, U.H., Dalal, M., Chatterji, M., Radha, D.R., Visweswariah, S.S., and Nagaraja, V. (2002) Functional characterisation of mycobacterial DNA gyrase: an efficient decatenase. *Nucleic Acids Res* **30**: 2144-2153.

Mansour, F.H., and Pestov, D.G. (2013) Separation of long RNA by agarose-formaldehyde gel electrophoresis. *Anal Biochem* **441**: 18-20.

Martínez, L.C., Banda, M.M., Fernández-Mora, M., Santana, F.J., and Bustamante, V.H. (2014) HilD induces expression of *Salmonella* pathogenicity island 2 genes by displacing the global negative regulator H-NS from *ssrAB*. *J Bacteriol* **196**: 3746-3755.

Mathy, N., Bénard, L., Pellegrini, O., Daou, R., Wen, T., and Condon, C. (2007) 5'-to-3' exoribonuclease activity in bacteria: role of RNase J1 in rRNA maturation and 5' stability of mRNA. *Cell* **129**: 681-692.

McCarty, N.S., and Ledesma-Amaro, R. (2019) Synthetic Biology Tools to Engineer Microbial Communities for Biotechnology. *Trends Biotechnol* **37**: 181-197.

McGary, K., and Nudler, E. (2013) RNA polymerase and the ribosome: the close relationship. *Curr Opin Microbiol* **16**: 112-117.

Mendell, J.E., Clements, K.D., Choat, J.H., and Angert, E.R. (2008) Extreme polyploidy in a large bacterium. *Proc Natl Acad Sci U S A* **105**: 6730-6734.

Menzel, R., and Gellert, M. (1987) Fusions of the *Escherichia coli gyrA* and *gyrB* control regions to the galactokinase gene are inducible by coumermycin treatment. *J Bacteriol* **169**: 1272-1278.

Mercier, R., Petit, M.A., Schbath, S., Robin, S., El Karoui, M., Boccard, F., and Espéli, O. (2008) The MatP/matS site-specific system organizes the terminus region of the *E. coli* chromosome into a macrodomain. *Cell* **135**: 475-485.

Mettert, E.L., and Kiley, P.J. (2007) Contributions of [4Fe-4S]-FNR and integration host factor to *fnr* transcriptional regulation. *J Bacteriol* **189**: 3036-3043.

Meyer, A.S., and Grainger, D.C. (2013) The *Escherichia coli* nucleoid in stationary phase. *Adv Appl Microbiol* **83**: 69-86.

Michalski, A., Damoc, E., Hauschild, J.P., Lange, O., Wieghaus, A., Makarov, A., Nagaraj, N., Cox, J., Mann, M., and Horning, S. (2011) Mass spectrometry-based proteomics using Q Exactive, a high-performance benchtop quadrupole Orbitrap mass spectrometer. *Mol Cell Proteomics* **10**: M111.011015.

Miller, H.I., and Friedman, D.I. (1980) An *E. coli* gene product required for lambda site-specific recombination. *Cell* **20**: 711-719.

Mirkin, E.V., and Mirkin, S.M. (2005) Mechanisms of transcription-replication collisions in bacteria. *Mol Cell Biol* **25**: 888-895.

Montero Llopis, P., Jackson, A.F., Sliusarenko, O., Surovtsev, I., Heinritz, J., Emonet, T., and Jacobs-Wagner, C. (2010) Spatial organization of the flow of genetic information in bacteria. *Nature* **466**: 77-81.

Monterroso, B., Zorrilla, S., Sobrinos-Sanguino, M., Robles-Ramos, M.A., López-Álvarez, M., Margolin, W., Keating, C.D., and Rivas, G. (2019) Bacterial FtsZ protein forms phase-separated condensates with its nucleoid-associated inhibitor SlmA. *EMBO Rep* **20**.

Morais Cabral, J.H., Jackson, A.P., Smith, C.V., Shikotra, N., Maxwell, A., and Liddington, R.C. (1997) Crystal structure of the breakage-reunion domain of DNA gyrase. *Nature* **388**: 903-906.

Murphy, K.C. (1998) Use of bacteriophage lambda recombination functions to promote gene replacement in *Escherichia coli*. *J Bacteriol* **180**: 2063-2071.
Murray, G.L., Attridge, S.R., and Morona, R. (2006) Altering the length of the lipopolysaccharide O antigen has an impact on the interaction of *Salmonella enterica* serovar Typhimurium with macrophages and complement. *J Bacteriol* **188**: 2735-2739.
Nanji, T., Gehrke, E.J., Shen, Y., Gloyd, M., Zhang, X., Firby, C.D., Huynh, A., Razi, A., Ortega, J., Elliot, M.A., and Guarné, A. (2019) *Streptomyces* IHF uses multiple interfaces to bind DNA. *Biochim Biophys Acta Gen Subj* **1863**: 129405.
Nelson, D.L., and Kennedy, E.P. (1971) Magnesium transport in *Escherichia coli*. Inhibition by cobaltous ion. *J Biol Chem* **246**: 3042-3049.
Neumann, S., and Quiñones, A. (1997) Discoordinate gene expression of gyrA and gyrB in response to DNA gyrase inhibition in *Escherichia coli*. *J Basic Microbiol* **37**: 53-69.
Nikaido, H. (2003) Molecular basis of bacterial outer membrane permeability revisited. *Microbiol Mol Biol Rev* **67**: 593-656.
Niki, H., Jaffé, A., Imamura, R., Ogura, T., and Hiraga, S. (1991) The new gene mukB codes for a 177 kd protein with coiled-coil domains involved in chromosome partitioning of E. coli. *Embo j* **10**: 183-193.
Nuñez, J.K., Bai, L., Harrington, L.B., Hinder, T.L., and Doudna, J.A. (2016) CRISPR immunological memory requires a host factor for specificity. *Mol Cell* **62**: 824-833.
O'Croinin, T., Carroll, R.K., Kelly, A., and Dorman, C.J. (2006) Roles for DNA supercoiling and the Fis protein in modulating expression of virulence genes during intracellular growth of *Salmonella enterica* serovar Typhimurium. *Mol Microbiol* **62**: 869-882.
Odermatt, N.T., Lelli, M., Herrmann, T., Abriata, L.A., Japaridze, A., Voilquin, H., Singh, R., Piton, J., Emsley, L., Dietler, G., and Cole, S.T. (2020) Structural and DNA binding properties of mycobacterial integration host factor mIHF. *J Struct Biol* **209**: 107434.
Osbourn, A.E., and Field, B. (2009) Operons. *Cell Mol Life Sci* **66**: 3755-3775.
Parry, C.M., Hien, T.T., Dougan, G., White, N.J., and Farrar, J.J. (2002) Typhoid fever. *N Engl J Med* **347**: 1770-1782.
Perez-Cheeks, B.A., Lee, C., Hayama, R., and Marians, K.J. (2012) A role for topoisomerase III in *Escherichia coli* chromosome segregation. *Mol Microbiol* **86**: 1007-1022.
Peterson, L.W., and Artis, D. (2014) Intestinal epithelial cells: regulators of barrier function and immune homeostasis. *Nat Rev Immunol* **14**: 141-153.
Porter, M.E., and Dorman, C.J. (1997) Positive regulation of *Shigella flexneri* virulence genes by integration host factor. *J Bacteriol* **179**: 6537-6550.
Postow, L., Hardy, C.D., Arsuaga, J., and Cozzarelli, N.R. (2004) Topological domain structure of the *Escherichia coli* chromosome. *Genes Dev* **18**: 1766-1779.
Potrykus, K., and Cashel, M. (2008) (p)ppGpp: still magical? *Annu Rev Microbiol* **62**: 35-51.
Pratt, T.S., Steiner, T., Feldman, L.S., Walker, K.A., and Osuna, R. (1997) Deletion analysis of the *fis* promoter region in *Escherichia coli*: antagonistic effects of integration host factor and Fis. *J Bacteriol* **179**: 6367-6377.
Pretorius, I.S., and Boeke, J.D. (2018) Yeast 2.0-connecting the dots in the construction of the world's first functional synthetic eukaryotic genome. *FEMS Yeast Res* **18**.
Price, M.N., Arkin, A.P., and Alm, E.J. (2006) The life-cycle of operons. *PLoS Genet* **2**: e96.
Prieto, A.I., Kahramanoglou, C., Ali, R.M., Fraser, G.M., Seshasayee, A.S., and Luscombe, N.M. (2012) Genomic analysis of DNA binding and gene regulation by homologous nucleoid-associated proteins IHF and HU in *Escherichia coli* K12. *Nucleic Acids Res* **40**: 3524-3537.
Pruss, G.J., Manes, S.H., and Drlica, K. (1982) *Escherichia coli* DNA topoisomerase I mutants: increased supercoiling is corrected by mutations near gyrase genes. *Cell* **31**: 35-42.

Qin, L., Erkelens, A.M., Ben Bdira, F., and Dame, R.T. (2019) The architects of bacterial DNA bridges: a structurally and functionally conserved family of proteins. *Open Biol* **9**: 190223.

Quinn, H.J., Cameron, A.D., and Dorman, C.J. (2014) Bacterial regulon evolution: distinct responses and roles for the identical OmpR proteins of *Salmonella* Typhimurium and *Escherichia coli* in the acid stress response. *PLoS Genet* **10**: e1004215.

Reuter, K., Sanderbrand, S., Jomaa, H., Wiesner, J., Steinbrecher, I., Beck, E., Hintz, M., Klebe, G., and Stubbs, M.T. (2002) Crystal structure of 1-deoxy-D-xylulose-5-phosphate reductoisomerase, a crucial enzyme in the non-mevalonate pathway of isoprenoid biosynthesis. *J Biol Chem* **277**: 5378-5384.

Reyes-Lamothe, R., Nicolas, E., and Sherratt, D.J. (2012) Chromosome replication and segregation in bacteria. *Annu Rev Genet* **46**: 121-143.

Reyes-Lamothe, R., Possoz, C., Danilova, O., and Sherratt, D.J. (2008) Independent positioning and action of *Escherichia coli* replisomes in live cells. *Cell* **133**: 90-102.

Rice, P.A., Yang, S., Mizuuchi, K., and Nash, H.A. (1996) Crystal structure of an IHF-DNA complex: a protein-induced DNA U-turn. *Cell* **87**: 1295-1306.

Rocha, E.P. (2008) The organization of the bacterial genome. *Annu Rev Genet* **42**: 211-233.

Rohs, R., West, S.M., Sosinsky, A., Liu, P., Mann, R.S., and Honig, B. (2009) The role of DNA shape in protein-DNA recognition. *Nature* **461**: 1248-1253.

Romeo, T., and Babitzke, P. (2018) Global regulation by CsrA and its RNA antagonists. *Microbiol Spectr* **6**.

Romeo, T., Gong, M., Liu, M.Y., and Brun-Zinkernagel, A.M. (1993) Identification and molecular characterization of csrA, a pleiotropic gene from *Escherichia coli* that affects glycogen biosynthesis, gluconeogenesis, cell size, and surface properties. *J Bacteriol* **175**: 4744-4755.

Rovinskiy, N.S., Agbleke, A.A., Chesnokova, O.N., and Higgins, N.P. (2019) Supercoil levels in *E. coli* and *Salmonella* chromosomes are regulated by the C-Terminal 35–38 amino acids of GyrA. *Microorganisms* **7**.

Ryan, V.T., Grimwade, J.E., Camara, J.E., Crooke, E., and Leonard, A.C. (2004) Escherichia coli prereplication complex assembly is regulated by dynamic interplay among Fis, IHF and DnaA. *Mol Microbiol* **51**: 1347-1359.

Saiki, R.K., Scharf, S., Faloona, F., Mullis, K.B., Horn, G.T., Erlich, H.A., and Arnheim, N. (1985) Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science* **230**: 1350-1354.

Samadpour, A.N., and Merrikh, H. (2018) DNA gyrase activity regulates DnaA-dependent replication initiation in *Bacillus subtilis*. *Mol Microbiol* **108**: 115-127.

Sambrook, J., and Russell, D.W. (2006) Purification of nucleic acids by extraction with phenol:chloroform. *CSH Protoc* **2006**.

Santero, E., Hoover, T.R., North, A.K., Berger, D.K., Porter, S.C., and Kustu, S. (1992) Role of integration host factor in stimulating transcription from the sigma 54-dependent *nifH* promoter. *J Mol Biol* **227**: 602-620.

Savery, N., Rhodius, V., and Busby, S. (1996) Protein-protein interactions during transcription activation: the case of the *Escherichia coli* cyclic AMP receptor protein. *Philos Trans R Soc Lond B Biol Sci* **351**: 543-550.

Sayers, E.W., Cavanaugh, M., Clark, K., Ostell, J., Pruitt, K.D., and Karsch-Mizrachi, I. (2019) GenBank. *Nucleic Acids Res* **47**: D94-D99.

Schavemaker, P.E., Boersma, A.J., and Poolman, B. (2018) How important is protein diffusion in prokaryotes? *Front Mol Biosci* **5**: 93.

Schmid, M.B., and Roth, J.R. (1987) Gene location affects expression level in *Salmonella* Typhimurium. *J Bacteriol* **169**: 2872-2875.

Schmidt, H., and Hensel, M. (2004) Pathogenicity islands in bacterial pathogenesis. *Clin Microbiol Rev* **17**: 14-56.

Schmieger, H. (1972) Phage P22-mutants with increased or decreased transduction abilities. *Mol Gen Genet* **119**: 75-88.

Schneider, R., Lurz, R., Lüder, G., Tolksdorf, C., Travers, A., and Muskhelishvili, G. (2001) An architectural role of the *Escherichia coli* chromatin protein FIS in organising DNA. *Nucleic Acids Res* **29**: 5107-5114.

Schneider, R., Travers, A., Kutateladze, T., and Muskhelishvili, G. (1999) A DNA architectural protein couples cellular physiology and DNA topology in *Escherichia coli*. *Mol Microbiol* **34**: 953-964.

Schulze-Gahmen, U., Rini, J.M., Arevalo, J., Stura, E.A., Kenten, J.H., and Wilson, I.A. (1988) Preliminary crystallographic data, primary sequence, and binding data for an anti-peptide Fab and its complex with a synthetic peptide from influenza virus hemagglutinin. *J Biol Chem* **263**: 17100-17105.

Schuwirth, B.S., Borovinskaya, M.A., Hau, C.W., Zhang, W., Vila-Sanjurjo, A., Holton, J.M., and Cate, J.H. (2005) Structures of the bacterial ribosome at 3.5 A resolution. *Science* **310**: 827-834.

Schwanhäusser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W., and Selbach, M. (2011) Global quantification of mammalian gene expression control. *Nature* **473**: 337-342.

Seah, N.E., Warren, D., Tong, W., Laxmikanthan, G., Van Duyne, G.D., and Landy, A. (2014) Nucleoprotein architectures regulating the directionality of viral integration and excision. *Proc Natl Acad Sci U S A* **111**: 12372-12377.

Shahul Hameed, U.F., Liao, C., Radhakrishnan, A.K., Huser, F., Aljedani, S.S., Zhao, X., Momin, A.A., Melo, F.A., Guo, X., Brooks, C., Li, Y., Cui, X., Gao, X., Ladbury, J.E., Jaremko, Ł., Jaremko, M., Li, J., and Arold, S.T. (2019) H-NS uses an autoinhibitory conformational switch for environment-controlled gene silencing. *Nucleic Acids Res* **47**: 2666-2680.

Sharadamma, N., Harshavardhana, Y., Ravishankar, A., Anand, P., Chandra, N., and Muniyappa, K., (2014) Molecular dissection of *Mycobacterium tuberculosis* integration host factor reveals novel insights into the mode of DNA binding and nucleoid compaction. *J Biol Chem.* The American Society for Biochemistry and Molecular Biology., pp. 34325-34340.

Shea, J.E., Hensel, M., Gleeson, C., and Holden, D.W. (1996) Identification of a virulence locus encoding a second type III secretion system in *Salmonella* Typhimurium. *Proc Natl Acad Sci U S A* **93**: 2593-2597.

Sheridan, S.D., Benham, C.J., and Hatfield, G.W. (1998) Activation of gene expression by a novel DNA structural transmission mechanism that requires supercoiling-induced DNA duplex destabilization in an upstream activating sequence. *J Biol Chem* **273**: 21298-21308.

Slager, J., and Veening, J.W. (2016) Hard-wired control of bacterial processes by chromosomal gene location. *Trends Microbiol* **24**: 788-800.

Sobetzko, P. (2016) Transcription-coupled DNA supercoiling dictates the chromosomal arrangement of bacterial genes. *Nucleic Acids Res* **44**: 1514-1524.

Sobetzko, P., Travers, A., and Muskhelishvili, G. (2012) Gene order and chromosome dynamics coordinate spatiotemporal gene expression during the bacterial growth cycle. *Proc Natl Acad Sci U S A* **109**: E42-50.

Soler-Bistué, A., Timmermans, M., and Mazel, D. (2017) The proximity of ribosomal protein genes to. *mBio* **8**.

Sridhar, S., and Steele-Mortimer, O. (2016) Inherent variability of growth media impacts the ability of *Salmonella* Typhimurium to interact with host cells. *PLoS One* **11**: e0157043.

Srikumar, S., Kröger, C., Hébrard, M., Colgan, A., Owen, S.V., Sivasankaran, S.K., Cameron, A.D., Hokamp, K., and Hinton, J.C. (2015) RNA-seq brings new insights to the intra-macrophage transcriptome of *Salmonella* Typhimurium. *PLoS Pathog* **11**: e1005262.
Srivatsan, A., Tehranchi, A., MacAlpine, D.M., and Wang, J.D. (2010) Co-orientation of replication and transcription preserves genome integrity. *PLoS Genet* **6**: e1000810.
Steck, T.R., and Drlica, K. (1984) Bacterial chromosome segregation: evidence for DNA gyrase involvement in decatenation. *Cell* **36**: 1081-1088.
Steele-Mortimer, O., Méresse, S., Gorvel, J.P., Toh, B.H., and Finlay, B.B. (1999) Biogenesis of *Salmonella* Typhimurium-containing vacuoles in epithelial cells involves interactions with the early endocytic pathway. *Cell Microbiol* **1**: 33-49.
Stoebel, D.M., Free, A., and Dorman, C.J. (2008) Anti-silencing: overcoming H-NS-mediated repression of transcription in Gram-negative enteric bacteria. *Microbiology (Reading)* **154**: 2533-2545.
Stracy, M., Wollman, A.J.M., Kaja, E., Gapinski, J., Lee, J.E., Leek, V.A., McKie, S.J., Mitchenall, L.A., Maxwell, A., Sherratt, D.J., Leake, M.C., and Zawadzki, P. (2019) Single-molecule imaging of DNA gyrase activity in living *Escherichia coli*. *Nucleic Acids Res* **47**: 210-220.
Sturm, A., Heinemann, M., Arnoldini, M., Benecke, A., Ackermann, M., Benz, M., Dormann, J., and Hardt, W.D. (2011) The cost of virulence: retarded growth of *Salmonella* Typhimurium cells expressing type III secretion system 1. *PLoS Pathog* **7**: e1002143.
Sutormin, D., Rubanova, N., Logacheva, M., Ghilarov, D., and Severinov, K. (2019) Single-nucleotide-resolution mapping of DNA gyrase cleavage sites across the *Escherichia coli* genome. *Nucleic Acids Res* **47**: 1601.
Swiercz, J.P., Nanji, T., Gloyd, M., Guarné, A., and Elliot, M.A. (2013) A novel nucleoid-associated protein specific to the actinobacteria. *Nucleic Acids Res* **41**: 4171-4184.
Swinger, K.K., and Rice, P.A. (2004) IHF and HU: flexible architects of bent DNA. *Curr Opin Struct Biol* **14**: 28-35.
Szafran, M.J., Gongerowska, M., Gutkowski, P., Zakrzewska-Czerwińska, J., and Jakimowicz, D. (2016) The coordinated positive regulation of topoisomerase genes maintains topological homeostasis in *Streptomyces coelicolor*. *J Bacteriol* **198**: 3016-3028.
Szklarczyk, D., Franceschini, A., Wyder, S., Forslund, K., Heller, D., Huerta-Cepas, J., Simonovic, M., Roth, A., Santos, A., Tsafou, K.P., Kuhn, M., Bork, P., Jensen, L.J., and von Mering, C. (2015) STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res* **43**: D447-452.
Takahashi, S., Kuzuyama, T., Watanabe, H., and Seto, H. (1998) A 1-deoxy-D-xylulose 5-phosphate reductoisomerase catalyzing the formation of 2-C-methyl-D-erythritol 4-phosphate in an alternative nonmevalonate pathway for terpenoid biosynthesis. *Proc Natl Acad Sci U S A* **95**: 9879-9884.
Tigges, M., Marquez-Lago, T.T., Stelling, J., and Fussenegger, M. (2009) A tunable synthetic mammalian oscillator. *Nature* **457**: 309-312.
Tsai, H.H., Huang, C.H., Tessmer, I., Erie, D.A., and Chen, C.W. (2011) Linear *Streptomyces* plasmids form superhelical circles through interactions between their terminal proteins. *Nucleic Acids Res* **39**: 2165-2174.
Tsui, P., Huang, L., and Freundlich, M. (1991) Integration host factor binds specifically to multiple sites in the *ompB* promoter of *Escherichia coli* and inhibits transcription. *J Bacteriol* **173**: 5800-5807.
Tusher, V.G., Tibshirani, R., and Chu, G. (2001) Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci U S A* **98**: 5116-5121.
Tyanova, S., Temu, T., Sinitcyn, P., Carlson, A., Hein, M.Y., Geiger, T., Mann, M., and Cox, J. (2016) The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat Methods* **13**: 731-740.

Unniraman, S., and Nagaraja, V. (1999) Regulation of DNA gyrase operon in *Mycobacterium smegmatis*: a distinct mechanism of relaxation stimulated transcription. *Genes Cells* **4**: 697-706.

Uzzau, S., Figueroa-Bossi, N., Rubino, S., and Bossi, L. (2001) Epitope tagging of chromosomal genes in *Salmonella*. *Proc Natl Acad Sci U S A* **98**: 15264-15269.

Valens, M., Penaud, S., Rossignol, M., Cornet, F., and Boccard, F. (2004) Macrodomain organization of the *Escherichia coli* chromosome. *EMBO J* **23**: 4330-4341.

van der Valk, R.A., Vreede, J., Qin, L., Moolenaar, G.F., Hofmann, A., Goosen, N., and Dame, R.T. (2017) Mechanism of environmentally driven conformational changes that modulate H-NS DNA-bridging activity. *Elife* **6**.

van der Velden, A.W., Lindgren, S.W., Worley, M.J., and Heffron, F. (2000) *Salmonella* pathogenicity island 1-independent induction of apoptosis in infected macrophages by *Salmonella enterica* serotype Typhimurium. *Infect Immun* **68**: 5702-5709.

van Workum, M., van Dooren, S.J., Oldenburg, N., Molenaar, D., Jensen, P.R., Snoep, J.L., and Westerhoff, H.V. (1996) DNA supercoiling depends on the phosphorylation potential in *Escherichia coli*. *Mol Microbiol* **20**: 351-360.

Vanden Broeck, A., Lotz, C., Ortiz, J., and Lamour, V. (2019a) Cryo-EM structure of the *complete E. coli* DNA gyrase nucleoprotein complex. *Nat Commun* **10**: 4935.

Vanden Broeck, A., McEwen, A.G., Chebaro, Y., Potier, N., and Lamour, V. (2019b) Structural Basis for DNA Gyrase Interaction with Coumermycin A1. *J Med Chem* **62**: 4225-4231.

Velmurugu, Y., Vivas, P., Connolly, M., Kuznetsov, S.V., Rice, P.A., and Ansari, A. (2018) Two-step interrogation then recognition of DNA binding site by Integration Host Factor: an architectural DNA-bending protein. *Nucleic Acids Res* **46**: 1741-1755.

Volff, J.N., and Altenbuchner, J. (2000) A new beginning with new ends: linearisation of circular chromosomes during bacterial evolution. *FEMS Microbiol Lett* **186**: 143-150.

Volkmer, B., and Heinemann, M. (2011) Condition-dependent cell volume and concentration of *Escherichia coli* to facilitate data conversion for systems biology modeling. *PLoS One* **6**: e23126.

Vologodskii, A.V., Lukashin, A.V., Anshelevich, V.V., and Frank-Kamenetskii, M.D. (1979) Fluctuations in superhelical DNA. *Nucleic Acids Res* **6**: 967-982.

Waldminghaus, T., and Skarstad, K. (2009) The *Escherichia coli* SeqA protein. *Plasmid* **61**: 141-150.

Wang, M., Herrmann, C.J., Simonovic, M., Szklarczyk, D., and von Mering, C. (2015) Version 4.0 of PaxDb: Protein abundance data, integrated across model organisms, tissues, and cell-lines. *Proteomics* **15**: 3163-3168.

Wang, W., Li, G.W., Chen, C., Xie, X.S., and Zhuang, X. (2011) Chromosome organization by a nucleoid-associated protein in live bacteria. *Science* **333**: 1445-1449.

Weglenska, A., Jacob, B., and Sirko, A. (1996) Transcriptional pattern of *Escherichia coli ihfB* (*himD*) gene expression. *Gene* **181**: 85-88.

Weiner, J.H., and Li, L. (2008) Proteome of the *Escherichia coli* envelope and technological challenges in membrane proteome analysis. *Biochim Biophys Acta* **1778**: 1698-1713.

Wertheimer, S.J., Klotsky, R.A., and Schwartz, I. (1988) Transcriptional patterns for the *thrS-infC-rplT* operon of *Escherichia coli*. *Gene* **63**: 309-320.

Williams, N.L., and Maxwell, A. (1999) Probing the two-gate mechanism of DNA gyrase using cysteine cross-linking. *Biochemistry* **38**: 13502-13511.

Wood, D.W., Setubal, J.C., Kaul, R., Monks, D.E., Kitajima, J.P., Okura, V.K., Zhou, Y., Chen, L., Wood, G.E., Almeida, N.F., Jr., Woo, L., Chen, Y., Paulsen, I.T., Eisen, J.A., Karp, P.D., Bovee, D., Sr., Chapman, P., Clendenning, J., Deatherage, G., Gillet, W., Grant, C., Kutyavin, T., Levy, R., Li, M.J., McClelland, E., Palmieri, A., Raymond, C., Rouse, G., Saenphimmachak, C., Wu, Z., Romero, P., Gordon, D., Zhang, S., Yoo, H., Tao, Y., Biddle, P., Jung, M., Krespan, W., Perry, M., Gordon-Kamm, B., Liao, L., Kim, S.,

Hendrick, C., Zhao, Z.Y., Dolan, M., Chumley, F., Tingey, S.V., Tomb, J.F., Gordon, M.P., Olson, M.V., and Nester, E.W. (2001) The genome of the natural genetic engineer *Agrobacterium tumefaciens* C58. *Science* **294**: 2317-2323.

Yanisch-Perron, C., Vieira, J., and Messing, J. (1985) Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **33**: 103-119.

Yosef, I., Goren, M.G., and Qimron, U. (2012) Proteins and DNA elements essential for the CRISPR adaptation process in *Escherichia coli*. *Nucleic Acids Res* **40**: 5569-5576.

Yoshua, S.B., Watson, G.D., Howard, J.A.L., Velasco-Berrelleza, V., Leake, M.C., and Noy, A. (2020) A nucleoid-associated protein bends and bridges DNA in a multiplicity of topological states with varying specificity. *bioRxiv*: 2020.2004.2017.047076.

Yubero, P., and Poyatos, J.F. (2020) The Impact of Global Transcriptional Regulation on Bacterial Gene Order. *iScience* **23**: 101029.

Zerbino, D.R. (2010) Using the Velvet de novo assembler for short-read sequencing technologies. *Curr Protoc Bioinformatics* **Chapter 11**: Unit 11.15.

Zhang, A., Rimsky, S., Reaban, M.E., Buc, H., and Belfort, M. (1996) Escherichia coli protein analogs StpA and H-NS: regulatory loops, similar and disparate effects on nucleic acid dynamics. *Embo j* **15**: 1340-1349.

Zulianello, L., de la Gorgue de Rosny, E., van Ulsen, P., van de Putte, P., and Goosen, N. (1994) The HimA and HimD subunits of integration host factor can specifically bind to DNA as homodimers. *Embo j* **13**: 1534-1540.